


Summer 8-13-2018

# Understanding Student Development Of Science Literacy Skills in an Undergraduate Environmental Science Course

Molly Picillo

University of Maine, [molly.picillo@maine.edu](mailto:molly.picillo@maine.edu)

Follow this and additional works at: <https://digitalcommons.library.umaine.edu/etd>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), [Educational Methods Commons](#), [Environmental Education Commons](#), [Science and Mathematics Education Commons](#), and the [Secondary Education Commons](#)

---

## Recommended Citation

Picillo, Molly, "Understanding Student Development Of Science Literacy Skills in an Undergraduate Environmental Science Course" (2018). *Electronic Theses and Dissertations*. 2906.  
<https://digitalcommons.library.umaine.edu/etd/2906>

This Open-Access Thesis is brought to you for free and open access by DigitalCommons@UMaine. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of DigitalCommons@UMaine. For more information, please contact [um.library.technical.services@maine.edu](mailto:um.library.technical.services@maine.edu).

**UNDERSTANDING STUDENT DEVELOPMENT OF SCIENCE**

**LITERACY SKILLS IN AN UNDERGRADUATE**

**ENVIRONMENTAL SCIENCE COURSE**

By

Molly Picillo

B.S. University of Maine, 2016

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Teaching)

The Graduate School

The University of Maine

August 2018

Advisory Committee:

Sarah Nelson, Associate Research Professor, School of Forest Resources, and Director, Program  
in Ecology and Environmental Sciences, Co-Advisor

Katharine Ruskin, Lecturer and Undergraduate Coordinator in Ecology and Environmental  
Sciences, Co-Advisor

Janet Fairman, Associate Professor of Education and Co-director of the Maine Education Policy  
Research Institute

© 2018 Molly J. Picillo

All Rights Reserved

**UNDERSTANDING STUDENT DEVELOPMENT OF SCIENCE**

**LITERACY SKILLS IN AN UNDERGRADUATE**

**ENVIRONMENTAL SCIENCE COURSE**

By Molly Picillo

Thesis Advisors: Dr. Sarah Nelson and Dr. Katharine Ruskin

An Abstract of the Thesis Presented

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Teaching

August 2018

Educators are expected to instill a variety of skills in their students that are necessary to be competent citizens of society. One such set of skills, science literacy skills, broadly encompass the ability of an individual to evaluate reliability of data and information and critically analyze and interpret them (Gormally Brickman, Hallar, & Armstrong, 2009). These skills are utilized in everyday decision-making and given their pertinence, there is a need for citizens to be scientifically literate. Thus, educators need tools and assessments to help students develop these skills and analyze their science literacy. The aim of this study was to develop science literacy interventions that could be easily incorporated into college curricula, providing instructors with exemplars of classroom interventions with the intent to improve students' science literacy skills. Therefore, the broad research question for this investigation was: How do science literacy interventions impact student proficiency in science literacy skills in college general education courses? I measured effectiveness of the interventions using the Test of Science Literacy Skills (TOSLS, Gormally Brickman, & Lutz, 2012) pre- and post-survey scores, as well as student feedback from pre- and post-survey follow-up interviews. The TOSLS surveys were given as part of a participation grade to students in a general education undergraduate college course (n = 148). A subset of students

volunteered to be interviewed regarding specific questions from the TOSLS survey after both the pre-survey ( $n = 12$ ) and the post-survey ( $n = 5$ ) to further investigate student understanding and interpretation. Interventions were designed by modifying previous assignments from earlier years' offerings of the class and were conducted both during class and outside of class as homework extensions. These interventions were created by evaluating scores and interviews on the TOSLS survey deployed as a pilot study in a previous semester of the undergraduate course. Based on these pilot data, four survey questions encompassing different science literacy skills of particular difficulty were targeted for intervention. The interventions were: (1) an interactive clicker-based lesson involving graph selection methods (2) data summits involving graph interpretation and source evaluation and (3) a role-play after which students discussed sources of bias.

Although the results indicated no statistically significant changes in the average scores between the pre-survey and post-survey ( $t$  test,  $p = 0.82$ ,  $\alpha = 0.05$ ), interviewed students recalled participating in the interventions and found them useful. Pre-survey scores ranged from 18%-96% correct with a mean score of 59%. Post-surveys had a slightly smaller range of 21%-96% with a mean of 60% correct. Based on these results, more work is necessary to provide instructors with course interventions that incorporate science literacy activities that target specific components of science literacy skills. Assessments, like TOSLS, are tools that can measure science literacy skills broadly across various science courses and provide a good overview of student science literacy. By broadening the use of a single tool, measurements can be compared between classrooms to produce interventions that do not have to heavily impact curriculum pacing, yet will provide students with the tools and skills necessary to be more scientifically literate citizens.

## DEDICATION

I would like to dedicate this thesis to **Luke Picillo** who has always inspired my passion to extend and deepen my understanding of science. Here is to boundless education and exploration. “Welcome to the life of a scientist!”

## **ACKNOWLEDGEMENTS**

I would first like to thank my thesis advisor Dr. Sarah Nelson of the School of Forest Resources at the University of Maine. Dr. Nelson took me under her wing and helped me explore a realm of research- education research- which was foreign to both of us. She never let my endless questioning discourage her faith in me, instead she aided me at every step. She helped me to keep strict deadlines and checked in to ensure my progress was where it should be and my spirits were kept up.

I would also like to thank my co-advisor Dr. Katharine Ruskin of the Ecology and Environmental Sciences program and Dr. Janet Fairman of the College of Education and Human Development both at the University of Maine. Dr. Ruskin and Dr. Fairman provided invaluable feedback throughout the research and writing processes. As primary instructor of the EES 100 course involved in this study, Dr. Ruskin helped greatly with the design and modifications of the interventions. Dr. Fairman's background in education research helped aid in the analysis of the data collected for this study.

I would like to thank the Research in STEM Education (RiSE) Center for providing me with both the resources necessary to carry out this thesis and connections to the individuals involved with this thesis. Individual thanks go to both Erin Vinson and Laura Millay who provided me with unyielding help and support. Without these two women, I would have been a sinking ship navigating through this new field of research. I also extend my thanks to the University of Maine for housing such a unique program as the RiSE Center and for hosting the students and the course involved in this thesis.

Finally, I must express my very profound gratitude to my family and friends who helped in the pilot tasks and also provided me with continuous support and encouragement throughout my life and completion of this thesis. This accomplishment would not have been possible without them.

Thank you.

Molly Picillo

## TABLE OF CONTENTS

DEDICATION .....	iii
ACKNOWLEDGEMENTS .....	iv
LIST OF TABLES .....	vii
LIST OF FIGURES .....	ix

### **CHAPTERS**

1. INTRODUCTION .....	1
1.1. Understanding of Science Literacy .....	1
1.2. Description of Study .....	2
2. LITERATURE REVIEW .....	4
2.1. Defining the Concept .....	4
2.2. The Need for Science Literacy .....	5
2.3. Tools for Assessing Science Literacy .....	7
2.4. Summary of Science Literacy Assessment .....	11
2.5. Teaching Science Literacy .....	11
3. METHODS .....	14
3.1. Context for the Study .....	14
3.2. Pilot Study .....	15
3.3. Study Structure .....	17
3.3.1. Pre-Survey and Post-Survey Methods .....	17
3.3.2. Pre-Interview and Post-Interview Methods .....	18
3.3.3. Intervention 1: Graph Choice Chart .....	19
3.3.4. Intervention 2: Data Summits .....	23
3.3.5. Intervention 3: Stakeholders in Research Role Play .....	25



3.4. Data Quality Assurance (QA), Preparation, and Statistical Analyses .....	25
4. RESULTS .....	29
4.1. Measuring Improvement Over the Course .....	29
4.2. Source Evaluation .....	32
4.2.1. Question 17 .....	33
4.2.2. Question 22 .....	34
4.3. Graphical Representation and Interpretation .....	37
4.3.1. Question 15 .....	38
4.3.2. Question 28 .....	39
5. DISCUSSION .....	42
5.1. Conclusions of the Study .....	42
5.2. Limitations and Modifications .....	45
5.3. Implications .....	46
REFERENCES .....	48
APPENDIX A: TEST OF SCIENCE LITERACY SKILLS SURVEY .....	51
APPENDIX B: PILOT INTERVIEW SCRIPT .....	62
APPENDIX C: PRE-INTERVIEW SCRIPT .....	63
APPENDIX D: POST-INTERVIEW SCRIPT .....	64
APPENDIX E: DATA SUMMIT- “IS THE ENVIRONMENT OF MAINE PRISTINE?” .....	66
APPENDIX F: DATA SUMMIT- “ARE WE IN THE ANTHROPOCENE?” .....	67
APPENDIX G: ROLE PLAY .....	68
APPENDIX H: CLASS DATA (MATCHED) .....	70
APPENDIX I: CLASS DATA (UNMATCHED) .....	71
BIOGRAPHY OF THE AUTHOR .....	72

## LIST OF TABLES

Table 1. Test of Science Literacy Skills (TOSLS) target question skills and text for EES 100 interviews, Spring 2018, University of Maine .....	16
Table 2. Timeline of events over the Spring 2018 semester of EES 100, University of Maine .....	17
Table 3. Interviewee demographics for the subset of EES 100 students who volunteered for interviews, Spring 2018, University of Maine .....	19
Table 4. Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between men and women .....	30
Table 5. Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between EES and non-EES majors .....	30
Table 6. Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between first-years and upperclass students .....	31
Table 7. Test of Science Literacy Skills (TOSLS) question analysis McNemar results .....	32
Table 8. Test of Science Literacy Skills (TOSLS) question analysis intervention participation Chi-squared results .....	32
Table 9. Test of Science Literacy Skills (TOSLS) Q17 McNemar’s contingency table in EES 100, Spring 2018, University of Maine .....	34
Table 10. Test of Science Literacy Skills (TOSLS) Q22 McNemar’s contingency table in EES 100, Spring 2018, University of Maine .....	35
Table 11. Interviewee use of the term “bias” or “unbiased” in Test of Science Literacy Skills (TOSLS) Q17 and Q22 for EES 100, Spring 2018, University of Maine .....	37
Table 12. Test of Science Literacy Skills (TOSLS) Q15 McNemar’s contingency table in EES 100, Spring 2018, University of Maine .....	38

Table 13. Test of Science Literacy Skills (TOSLS) Q28 McNemar’s contingency table in EES 100,

Spring 2018, University of Maine ..... 40

## LIST OF FIGURES

Figure 1. Graph Choice Chart .....	21
Figure 2. Test of Science Literacy Skills (TOSLS) pre- and post-survey scores for Q17 and Q22 in EES 100, Spring 2018, University of Maine .....	33
Figure 3. Test of Science Literacy Skills (TOSLS) Q17 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'B') .....	34
Figure 4. Test of Science Literacy Skills (TOSLS) Q22 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'C') .....	36
Figure 5. Test of Science Literacy Skills (TOSLS) pre- and post-survey scores for Q15 and Q28 in EES 100, Spring 2018, University of Maine .....	37
Figure 6. Test of Science Literacy Skills (TOSLS) Q15 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'D') .....	39
Figure 7. Test of Science Literacy Skills (TOSLS) Q28 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'B') .....	41

## CHAPTER 1

### INTRODUCTION

#### 1.1 Understanding of Science Literacy

In day-to-day life we are all expected to collect information from our surroundings, synthesize what we have experienced, and determine the appropriate action for that situation. This process, although seemingly complex, is the expectation for sound decision-making. Without the tools to analyze information, it can be hard to make an informed decision. Determining what to wear outside based on a weather report, choosing what product to buy given information on quality and price, or looking into research supporting medical procedures and medications are tasks that require data analysis. It is important to critically interpret new information and combine this with more extensive research or previously gained knowledge. Someone who can synthesize these different trusted sources of data logically into a decision is considered to be scientifically literate (Schielke, 2013). However, science literacy is a term that varies somewhat in definition in the literature and also takes into account the ability to use trustworthy science to support results, decisions, and inferences (DeBoer, 2000; Gormally et al., 2012; Koltay, 2016; Majima, 2015; Millar, 2006; and Morgan, Bertera, & Reid, 2007). This ability can incorporate interpreting a variety of quantitative, graphical, or statistical findings in making different kinds of decisions in daily life and further extends into our economic, healthcare, and political decisions. Science literacy also involves being able to critically evaluate the sources from which the data originate. People need to be primed for making life decisions based on scientific information from reliable sources, a skill set that can be gained through formal education.

Educators are increasingly encouraged to provide students with science literacy skills to better equip students to make decisions as citizens. Feinstein (2010) critically evaluated what is known about the use of science literacy skills by outsiders to the field and suggested that people do not view things in daily life as a scientist, but take bits and pieces of their scientific knowledge and apply it to a situation.

The key characteristic of the scientifically literate individual that Feinstein (2010) referred to as a “competent outsider” is the previous knowledge that can be repurposed. Falk et al. (2016) evaluated adult education at science centers— vaguely described as public institutions where a range of science content is available with no specific means of education (e.g., tours, self-investigation, or classes) — and found a correlation between the level of education (along with wealth) of each adult and their presence at such centers. Although the impact of education incorporates other confounding factors and could have skewed their results, Falk et al. (2016) demonstrated that individuals with higher levels of education were more likely to have continued engagement with science after leaving formal education settings. The benefit of this continued engagement is the increase of knowledge or scientific awareness from which a “competent outsider” can derive scientific knowledge to apply to life situations. In addition to informal educational approaches to science education, formal science education approaches have been developed to aid in science literacy skill development. Inquiry-based labs were one type of classroom methodology shown to result in gains (approximately 2–4%) in science literacy skills (Gormally et al., 2009). Evaluation showed that these gains were linked with the use of science literacy skills in a way that an average citizen would use them. This study intends to continue such inquiry in the classroom setting and to develop more tools for the classroom to provide students with a wealth of science literacy skills and knowledge.

## **1.2 Description of Study**

The Test of Science Literacy Skills (TOSLS) survey is an assessment used in 2017 and 2018 in the University of Maine’s Human Population and the Global Environment (EES 100) general education course. This assessment is used to identify areas of difficulty that students are having with science literacy and track changes in their science literacy understanding after having completed the course (Gormally et al., 2012). Based on preliminary TOSLS data collected in a pilot year in Spring 2017, I was able to target skills of particular difficulty for students and interview some of those students to

understand the reasons they struggled with those questions. From this initial study, I targeted three specific science literacy skills derived from the TOSLS survey: creating graphical representations of data; evaluating validity of sources; and justifying inferences, predictions, and conclusions based on quantitative data. In order to help students develop these skills, I used interventions that specifically targeted these three skill sets during the course in Spring 2018 with the understanding that a one-time intervention is not always enough to change student skills.

This project aimed to provide initial trials of targeted interventions. Although Gormally et al. (2012) provided a well-validated instrument for assessing science literacy, the authors stopped short of recommending interventions to address students' lack of skills. Gormally et al. (2012) conveyed in their study that not all students, or every class, improved significantly over one semester. However, I propose that by gaining an understanding of what students struggle with coming into the class, then targeting those science literacy skills specifically with interventions and reiterating the importance of these skills throughout the semester, that students would improve their science literacy skills by the end of the semester. The post-survey (end-of-semester) interviews were intended to allow students to reference these interventions as they discussed how their approach to science literacy skills had changed.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1. Defining the Concept

Science literacy, a term that is used widely today, lacks a consistent definition across all fields of science, technology, engineering and math (STEM) education. This term, first coined in 1958 (Feinstein, 2010), has been much debated in order to refine the definition of what it means to be scientifically literate. The term is extensive and encompasses smaller divisions of scientific understanding such as graphical interpretation, source evaluation, and data analysis (Gormally et al., 2009). Thus, in order to fully understand the concept, it is important to investigate different definitions and interpretations of science literacy used by scientists and science educators as well as the suite of skills that comprise it. DeBoer (2000) explained how science literacy is the end goal of many science courses; however, not everyone agrees on what it means and therefore it is not taught uniformly. Despite the unclear definition, Millar (2006) aimed to improve the use of science literacy in secondary education. The authors loosely defined science literacy as an “understanding of the nature of scientific knowledge, of the ways in which it is obtained, checked, and refined, and of the characteristics of sound arguments in support of a claim or conclusion about some aspect of the natural world” (Millar, 2006).

Koltay (2016) suggested that a more effective view of science literacy is as a continuum and argued that science literacy is really built on a scaffold of literacy including data literacy, information literacy, and math literacy. Information literacy is the ability to identify when data are necessary, finding and accessing resources that contain those data, determining if those resources are reliable, and applying the results to a broader picture (Koltay, 2016). Similarly, data literacy involves understanding those reliable resources and the results, often in graphical or tabular form, with their inferred conclusions as well as evaluating if the methodology is sound and the scientific process is held to a standard of rigor (Koltay, 2016; Webber, Nelson, Weatherbee, Zoellick, & Schauffler, 2014). Math



literacy not only touches on following math rules but also applying those strategies in a way that is functional in daily life (Kihara & Witzel, 2014). One study that tried to incorporate all of these different sub-categories broke down science literacy into nine key skills with the overarching definition that science literacy is “recognizing and analyzing the use of methods of inquiry that lead to scientific knowledge and the ability to organize, analyze, and interpret quantitative data and scientific information” for use in “real-world situations beyond the classroom” (Gormally et al., 2012). The assessment Gormally et al. developed in 2012 – Test of Science Literacy Skills (TOSLS) – predates Koltay’s (2016) description of the continuum, yet it encompasses those ideas by organizing student knowledge across a spectrum of science literacy categories. For the purpose of this study, I will use the Gormally et al. (2012) definition given that it focuses on the importance of science literacy in realms beyond just the field of science; in particular, focusing on the use of scientific information in everyday life.

## **2.2. The Need for Science Literacy**

Science literacy plays a role in our society that is fundamental to being an educated citizen; thus, formalizing the teaching of these skills is vital for their fluidity of use throughout society. The day-to-day role that science literacy plays can be anything from understanding risks and benefits of various medications to understanding how well appliances perform based on evidence (Majima, 2015). Ideally, people should be equipped with the ability to observe, ask questions, gather evidence, and draw conclusions in order to make important and educated decisions (Morgan et al., 2007). Many of life’s important decisions rely on these skills and people often work through parts of this process without even thinking about them as data-related skills. However, research has shown that lay people do not use science literacy skills when faced with decision-making that would benefit from scientific thought such as political decisions or financial choices (Feinstein, 2010). Providing practice with these skills could allow people to make better-informed decisions and recognize key missing components in their daily evaluations of life situations.

Although important for day-to-day life, there are times when science literacy skills become of particular importance for decision making that influences many people and their futures. Matschullat (2015) used an example that explained the burden that can fall on taxpayers when environmental policy is modified without informed decision-making based on scientific information; there are implications for ecosystems, towns, cities, states, nations, and the Earth as a whole. Given their scale of impact, science literacy skills warrant inclusion in classroom curricula. However, science literacy is far more than a technique for the science classroom; rather, it prepares individuals for all types of judgments and negotiations that occur across disciplines (Matschullat, 2015). Nevertheless, understanding science and technology is not always recognized as useful by the citizens who unknowingly utilize it. Valdecasas and Correias (2010) evaluated a survey of the Spanish population from 2008, which found that only 9.6% of the population showed interest in science and technology versus the 28% that showed interest in medicine and 26.1% in sports. As scientific concerns in society are becoming more pressing, providing citizens with science literacy approaches and skills will help them understand the use of science literacy as well as aid in solving real world and potentially large-scale problems.

We are currently in the Information Era, where information and data are easily accessible. However, without proper teaching of science literacy skills, people may incorrectly self-educate, which could lead to misconceptions. For instance, Falk et al. (2016) surveyed 1,018 adults in Los Angeles to determine how they were obtaining scientific information and found that they chose to use books, magazines, internet, documentaries and videos to educate themselves about science. Although all these informal science education resources and conduits have the potential to provide people with accurate scientific information individuals can be misled with false information without skills for determining the reliability of sources. People need to understand the process behind obtaining scientific findings in order to distinguish between fact and propaganda. This is becoming increasingly important in a time when news reports are presenting conflicting claims, based on differing evidence, to the general public.

Therefore, it can be useful to provide a setting such as the classroom to provide people with the methods to utilize such informal science educational resources effectively.

### **2.3. Tools for Assessing Science Literacy**

The first step in preparing students to be scientifically literate citizens is to assess the extent to which the students are scientifically literate. An instructor must gather information on the difficulties students have with certain skills in order to identify the gaps in skills and to target them within classroom instruction. One particularly intensive dual-response instrument that assesses science literacy is Student Understanding of Science and Scientific Inquiry (SUSSI, Liang et al., 2008). This tool was used to evaluate 209 pre-service elementary education majors (some dual majoring in special education) using questions involving the nature of science and touched upon important science literacy topics like scientific methods and inferences (Liang et al., 2008). In this assessment, participants were asked to indicate the degree to which they agreed or disagreed with scientific statements. For example one statement was, “Scientists’ observations of the same event will be the same because scientists are objective” and the response options ranged from “strongly disagree” to “strongly agree” with the option to remain neutral (Liang et al., 2008). Responses to this assessment were compared to the opinions of expert scientists. Liang et al. (2008) reported that none of the participants had informed views (understand differing views and interpretations based on previous knowledge, perspective, and belief) on either the Likert Scale (strongly agree - strongly disagree) or constructed responses for scientific theory and laws. Further, only 14% of participants had informed views in each type of response related to scientific methods (Liang et al. 2008). Liang et al. (2008) showed that individuals answered questions of the same nature similarly; discrepancies were only found with the wording of open ended questions.

SUSSI can be used to guide instructional decisions or as a summative assessment for student achievement. This assessment strongly focuses on the data literacy component with less emphasis on information literacy. Therefore, the tool is not efficient for assessing the full scope of science literacy.

SUSSI is further limited by the narrow interpretation that can be gathered from assessments that rely on Likert Scales and the difficulty of drawing valid conclusions based on interpretation of variable written responses.

Blank et al. (2016) developed a novel, but unnamed tool designed for an in-depth evaluation of the information literacy skills of biology undergraduates. This assessment was administered over several homework assignments in which students were asked to find scientific literature that was relevant to a specified topic. Students (first year:  $n = 145$ ; seniors:  $n = 43$ ) were graded based on a rubric for source quality, source relevance, and citation quality. Blank et al. (2016) reported that 68% of first year students scored below a 67% on quality of source, whereas all seniors scored 67% or better (58% of seniors scored 100%). About 87% of students across both grade levels were able to find relevant sources and senior students did statistically significantly better than first year students in finding relevant peer-reviewed articles ( $p = 0.017$ ) and creating appropriate citations ( $p < 0.0001$ ). Alone, this resource is specific only to information literacy but if used in combination with SUSSI, this novel tool could provide a well-rounded assessment for science literacy. However, this combination of resources could be time-intensive and require a new rubric for assessing results.

Lawson's Classroom Test of Scientific Reasoning (LCTSR; Lawson, 1978) is a test that is one of the older assessment tools used for measuring science literacy. In this 75-100 minute classroom assessment, students watched a physical demonstration and then were asked to respond to a multiple-choice answer in a booklet where they had to follow up the answer with a written description of their reasoning behind their answer (Lawson, 1978). Correct answers were classified as having both the correct response to the multiple-choice question as well as correct reasoning in the written explanation. Results from the LCTSR surveys, which were distributed in required general science, biology, and English classes across one junior high and two high schools in the San Francisco Bay area ( $n = 513$  students), showed an increase in average scores from eighth through tenth grade; however, the mean score did

not exceed 60% at any grade level (Lawson, 1978). As Lawson (1978) concluded, this assessment helps instructors get an understanding of how students think which can help prepare lesson targets. Although this test is often referenced in science literacy papers, the content only incorporates the reasoning and data interpretation areas of science literacy and neglects the skill area of information literacy.

In reviewing the available, researched tools for measuring science literacy skills, it is clear that there is no one perfect tool to assess the broad suite of skills and concepts involved in science literacy. However, it is important to show why the assessment tool for this study was not arbitrary, rather it was the most suitable assessment. To maintain consistency across science literacy assessment in a way so that results obtained are comparable, a results single tool must be utilized and thus be widely accessible (for a relatively comprehensive list of assessments broken down by skill set, see Blank et al. (2016)). As such, the assessment selected for this study was the Test of Science Literacy Skills (TOSLS) survey (Gormally et al., 2012). TOSLS is useful in determining which skills students in large general education science classes have mastered and which skills need work (Gormally et al., 2012), and thus provides an appropriate tool for use at university (and smaller) scale. In particular, this survey is easy to administer given that it is accessible online to students and multiple choice, allowing for quick score turn-around.

This tool allocates science literacy into nine skills across two categories that are tested by answering 28 multiple-choice questions. The first category is “Understand methods of inquiry that lead to scientific knowledge” which is further broken down into four skills linked to information literacy that focus on how information is obtained in science and whether the information presented is accurate, reliable, and used in an appropriate context. The second category is “Organize, analyze, and interpret quantitative data and scientific information”, which incorporates the other five skills that focus on data and math literacy and include numerical results of a study, how those results are shown visually (graphing), and whether the conclusions based on those results are justified. In developing the tool, the authors situated the questions in real-world contexts, as is important when developing a set of practical

science literacy skills for scientists and non-scientists alike. For example, one question prompts students to think about their confidence in the accuracy of a website from a Google search of a topic they heard about on the news (Gormally et al., 2012). By applying these real life scenarios, the instrument is appropriate for a broader range of individuals than science students. Gormally et al. (2012) described the validity and reliability evaluations of the TOSLS survey, which included student interviews, biology educator expert reviews, pilot testing, examination of psychometric properties, classroom testing of the finalized instrument in multiple classrooms, and validation of the tool with different biology courses. Reliability was explored in the pre- and post-surveys using the Kuder-Richardson test; the pre-survey scored 0.731 and the post-survey scored 0.748, which is within the acceptable range for internal consistency. The validation process was imperative for ensuring confidence and trustworthiness of a newly developed research tool.

While investigating published TOSLS survey results from other institutions, it became apparent that although many authors have cited Gormally et al. (2012) results for the TOSLS survey are seldom reported. Rather, the authors typically reference the paper for its potential to be utilized in other educational assessments. For instance, Dasgupta, Anderson, and Pelaez (2014) developed a Rubric for Experimental Design (RED), utilizing the understanding that undergraduates struggle with the concept of representative samples in experimental design as described in Gormally et al. (2012). In this case, Dagsupta et al. (2014) were able to develop a rubric for diagnosing student difficulties with experimental design, but never utilized the TOSLS survey. Another study borrowed Gormally et al.'s (2012) finding that undergraduates have difficulties with statistical reasoning skills as a foundation for their research development of the Statistical Reasoning in Biology Concept Inventory (SRBCI; Deane, Nomme, Jeffery, Pollock, & Birol, 2016). Benjamin et al. (2014) used Gormally et al. (2012) as support for valid and reliable undergraduate surveying methods in the development of SLSCP, a survey that connects academic preparedness with science literacy. These are just a few examples that show that researchers

are not publishing results from the use of the TOSLS survey or about how to improve the tool, but rather using the student difficulties identified in Gormally et al. (2012) as a springboard for other research or instrument development. Without consistency in the use of a tool to measure science literacy skills within and across institutions, it is hard to further validate the TOSLS tool. Without consistent measurements of results, it is hard to document change and improvement over time. This study intends to create some consistency by using TOSLS instead of developing a new tool.

#### **2.4. Summary of Science Literacy Assessment**

As an assessment tool, TOSLS is able to provide an understanding of the skills with which students struggle most but few studies have implemented interventions to improve students' science literacy skills and measure the impact of those interventions by assessment. Lawson (1978), as well as others, explained that assessment tools provide instructors with an understanding of how students think, which can help prepare lesson targets. With this information, instructors need to try interventions that target these specific areas of student difficulty in order to build a toolbox for science literacy skill education. Testing interventions is an iterative process that needs to undergo trials and modification. Upon entering a science course, students have certain convictions and understandings of science that are years in the making (Lederman, 2002). Sometimes these preconceptions do not align with current scientific research, and may not be able to be altered after only one intervention. After implementation, TOSLS can be used to evaluate if there are changes in student understanding (Gormally et al., 2012), which can lead to further development and modification of science literacy interventions.

#### **2.5. Teaching Science Literacy**

Science literacy education can occur in a variety of settings including informal education centers. Valdecasas and Correias (2010) speak to the importance of museums (particularly natural history museums), which provide access to science collections and scientific knowledge not easily accessible from different parts of the world. They propose approaches that natural history museums can take to

promote science literacy, which include offering new perspectives and awareness to carefully consider and challenge existing conclusions. Morgan et al. (2007) described another science education resource designed by the foundation Setting Priorities for Retirement Years (SPRY) called Science Across the Generations (SAG). This resource consisted of 20 modules tested by both children ( $n = 1,568$ ) and adults ( $n = 1,471$ ). In these modules, participants were educated in science via reading, talking, writing, and the using new words with informal science materials. The response to SAG by adult participants was encouraging as 75% had learned something new about science and 56% learned something new about scientific problem solving (Morgan et al., 2007). Children showed on average 5% increase in science knowledge from pre- to post-instruction. Results from this study also showed that participants were excited about learning these science topics and that they were interested in continuing to learn. Not only do these informal education (non- classroom-based) opportunities excite the public to learn more, but by providing access to science education within communities (e.g. schools, education centers, community programs.), citizens are prepared to make more scientifically informed decisions within society.

Classroom instruction is another method of reaching out to people to improve scientific literacy within society. This allows for direct instruction, targeting broader audience that is not always reached by optional-attendance in informal settings. Similar to the modular approach of some informal settings, teachers and professors can prepare interventions to focus on topics of interest. Reeves and Honig (2016) deployed data literacy interventions with elementary pre-service teachers ( $n = 64$ ) utilizing a pre-post survey assessment to assess whether the participants had gains in knowledge and skills, attitude towards data, and data literacy skills after completing the intervention. The intervention involved two days of three-hour sessions. The first day was spent entering raw data into Excel datasets that provided statistical outputs and also running assessments for quality and reliability. The second day of intervention was spent on analysis and interpretation of the data. Participants were found to have gains



only on certain aspects of analysis; specifically, data literacy measures were significantly greater for the post-test (Reeves & Honig, 2016). Not only was this study beneficial as it incorporated pre-service teachers, but it was an exemplary study for the use of an intervention with undergraduate students. As was shown by Reeves and Honig (2016), not all interventions prove fruitful in all areas of intent such as changes in participants' attitudes and beliefs about data. But there is always something that can be learned from null data (Eskinasi & Fokkema 2006) that will provide insight for further inquiry.

## **CHAPTER 3**

### **METHODS**

#### **3.1. Context for the Study**

The focal course of this study is called “Human Population and the Global Environment” (EES 100), which is offered at the University of Maine—a land-grant state university located in a predominantly rural region in the northeastern U.S. This class is designed to introduce concepts and principles for evaluating the contemporary global issues of population growth, natural resource management, and environmental protection (The University of Maine, 2016). In addition to providing content knowledge about this wide array of environmental systems and issues, another objective of this three-credit course is for students to develop skills to critically interpret the diverse types of information available about environmental issues. This course satisfies the University's Population and the Environment general education requirement. The course also aims to achieve some of the Association of American Colleges and Universities (AACU) Valid Assessment of Learning in Undergraduate Education (VALUE) outcomes. For this assessment, rubrics are distributed to professors to aid in general education performances. Specifically, EES 100 aims for students to gain proficiency in five information literacy outcomes and four quantitative literacy outcomes. This class also benefits from being co-taught by two professors, as well as having two teaching assistants, and two Maine Learning Assistants (undergraduate peers that have taken the course previously and aid in student interaction and provide extra help for the current students).

EES 100 has no prerequisites and as an undergraduate general education course it is open to any undergraduate student attending the University of Maine. Because it fulfills a general education requirement, students in this class come from a variety of backgrounds, study different majors, and have different expected graduation dates. However, this course is also an early requirement for a degree in Ecology and Environmental Sciences (EES), the department offering the course. The class size ranges

from 100-150 students from five different colleges within the university. The gender ratio differs from year to year, but Spring 2018 had approximately a 50/50 gender split. First year students are most common (approximately 60%) with a spread of upperclass students tapering off towards seniors. Although EES students make up approximately a third of the class, many students take this course from other majors to meet the general education requirement.

### **3.2. Pilot Study**

In Spring 2017, the Test of Science Literacy Skills was distributed to students ( $n = 130$ ) in EES 100 as part of an Institutional Review Board (IRB) approved project at the university. To determine what science literacy skills would be best to target for the study, I calculated summary statistics for individual questions and evaluated the survey content for relevance to the course in consultation with the instructors. In order to further understand the results of the survey, I collaborated with the two instructors to develop clinical interviews. I tested those protocols with student volunteers ( $n = 5$ ) at the end of the semester in Spring 2017 to uncover the difficulties students were experiencing while answering questions 2, 5, 17, 22, and 28—those identified as addressing target skills—from the TOSLS survey (Survey: Appendix A; Interview: Appendix B). Based on an understanding of these difficulties through interviews, targeted interventions were developed for the class.

Given the results from EES 100 Spring 2017 pre- and post- TOSLS surveys and post-survey interviews, the instructors and I determined that questions 15, 17, 22, and 28 were of most interest and relevance to the class as they were in line with the course learning objectives and were areas of expected difficulty based on previous results. The skills associated with the questions were as follows: (1) conduct an effective literature search of sources and distinguish between types of sources (Q17 and Q22); (2) make a graph (Q15); and (3) justify inferences, predictions, and conclusions based on quantitative data (Q28) (Gormally et al., 2012). These questions were the focus of the Spring 2018 interview protocol that was given after the pre-TOSLS survey as well as after the post-TOSLS survey.

Pilot year interviews were lengthy and the phrasing of some interview questions led to responses that were drawn out and did not always provide conclusive information to help with course improvement. Thus, revisions were made to the pilot interview script in order to obtain more clear and concise responses from the participants and to enable more accurate coding for analysis. Table 1 provides the text of the selected interview questions for the Spring 2018 study, and identifies the related science literacy skill for each question.

**Table 1.** Test of Science Literacy Skills (TOSLS) target question skills and text for EES 100 interviews, Spring 2018, University of Maine.

<b>Question (abbrev. hereafter)</b>	<b>Science Literacy Skill (Gormally et al., 2012)</b>	<b>Text of Question (TOSLS; Gormally et al, 2012)</b> (For associated figures and answer options see Appendix A)
Question 15 (Q15)	Make a graph	Researchers found that chronically stressed individuals have significantly higher blood pressure compared to individuals with little stress. Which graph would be most appropriate for displaying the mean (average) blood pressure scores for high-stress and low-stress groups of people?
Question 17 (Q17)	Conduct an effective literature search of sources and distinguish between types of sources	The most important factor influencing you to categorize a research article as trustworthy science is:
Question 22 (Q22)		Your doctor prescribed you a drug that is brand new. The drug has some significant side effects, so you do some research to determine the effectiveness of the new drug compared to similar drugs on the market. Which of the following sources would provide the most accurate information?
Question 28 (Q28)	Justify inferences, predictions, and conclusions based on quantitative data	Researchers interested in the relation between River Shrimp ( <i>Macrobrachium</i> ) abundance and pool site elevation, presented the data in the graph below. Interestingly, the researchers also noted that water pools tended to be shallower at higher elevations. Which of the following is a plausible hypothesis to explain the results presented in the graph?

### 3.3 Study Structure

The study occurred over a 14-week semester and included several components, some of which occurred inside of class and others outside of class with certain voluntary options as well as participatory requirements (Table 2). Given the variation in those requirements, sample sizes were different for each study component.

**Table 2.** Timeline of events over the Spring 2018 semester of EES 100, The University of Maine.

Semester Week	Event	Sample size (N)	Approx. Time (min) Lecture (L) Homework (HW)
1-3	Pre-Survey	111	15-75 (HW)
1-3	Pre-Interviews	12	10-25 (optional)
3	Graph Choice Lecture	118	20 (L)
7	Data Summit 1	95	30 (L), 60 (HW)
10	Role Play	98	30 (L)
13	Data Summit 2	118	30 (L), 60 (HW)
13-14	Post-Survey	90	15-75 (HW)
13-14	Post-Interview	5	10-25 (optional)

#### 3.3.1. Pre-Survey and Post-Survey Methods

The Test of Science Literacy Skills (TOSLS; Gormally et al., 2012) survey was administered via an online platform associated with the course to all of the students ( $n = 148$ ). Students had two and a half weeks to complete the survey and were awarded participation credit for submitting it. Full credit was given for the submission; the score was not a factor in the distribution of credit. The post-survey was distributed through the same online forum and open to students approximately two weeks prior to the end of the semester. This ensured that students could complete the survey for class participation credit while also limiting the amount of time remaining in the class to best encompass what was learned over

the course of the semester. As was the case in the pre-survey, full points were awarded for the submission of the survey and were not based on the score. Student data were de-identified prior to analysis.

### **3.3.2. Pre-Interview and Post-Interview Methods**

Unlike the survey, interviews were completely voluntary but also had no weight on students' EES 100 course grade. Instead of course credit, a \$10 incentive was used to entice volunteers to participate. The survey was advertised both in-class and on the online forum and students were asked to answer four of the TOSLS questions in person. IRB approval was acquired to record the interviews for later analysis. For every TOSLS question asked, volunteers were asked to explain their reasoning behind their answer (interview script: Appendix C) and, although they were asked to take the survey prior to being interviewed, they were not required to answer the same way as they had in the survey. In addition, on the post-interview, the students were asked about EES 100 and the interventions (interview script: Appendix D). Interview participants could perform in both the pre-interviews and post-interviews but they were not required to do so. Course instructors were not aware of the names of participating students. Table 3 shows demographic information collected from the volunteers who participated in the interviews and which interview(s) they participated in. One student who took both the pre-interview and post-interview had not completed the survey prior to the pre-interview which is indicated by an asterisk (\*) in the table.

**Table 3.** Interviewee demographics for the subset of EES 100 students who volunteered for interviews, Spring 2018, University of Maine.

Interviewee	Interview(s)	Gender	Year	Major
1	Pre	Female	Upperclass	Nursing
2	Pre	Male	First-year	New Media
3	Pre	Male	Upperclass	EES
4*	Pre, Post	Male	First-year	Biology/EES
5	Pre	Female	Upperclass	Biology
6	Pre	Female	Upperclass	English
7	Pre, Post	Female	Upperclass	History
8	Pre	Male	First-year	Mechanical Engineering
9	Pre	Female	First-year	EES/Wildlife Ecology
10	Pre	Female	First-year	Engineering Physics
11	Pre	Female	Upperclass	EES
12	Pre, Post	Female	Upperclass	EES
13	Post	Male	Upperclass	EES
14	Post	Female	Upperclass	EES

\*This participant did not take the pre-survey prior to interview

### 3.3.3. Intervention 1: Graph Choice Chart

The Graph Choice Chart (GCC) is a tool designed by The Maine Data Literacy Project with the intent to help students use data as evidence to defend claims and also identify appropriate graphs for answering research questions (Figure 1; Webber et al., 2014). Since data can be presented in many different formats (graphs of all types, figures, maps, etc.) interpretation of data by readers can be difficult (Weissgerber, Milic, Winham, & Garovic, 2015). Thinking critically about the presentation of information and what story the data tell are important skills and will prepare students for thinking about how they will present and interpret data. Webber et al. (2014) found when surveying 200 high school students only 23% properly drew a graph to show the effect of environment on population abundance and 58% of students drew an appropriate graph to show correlation between two factors. Based on

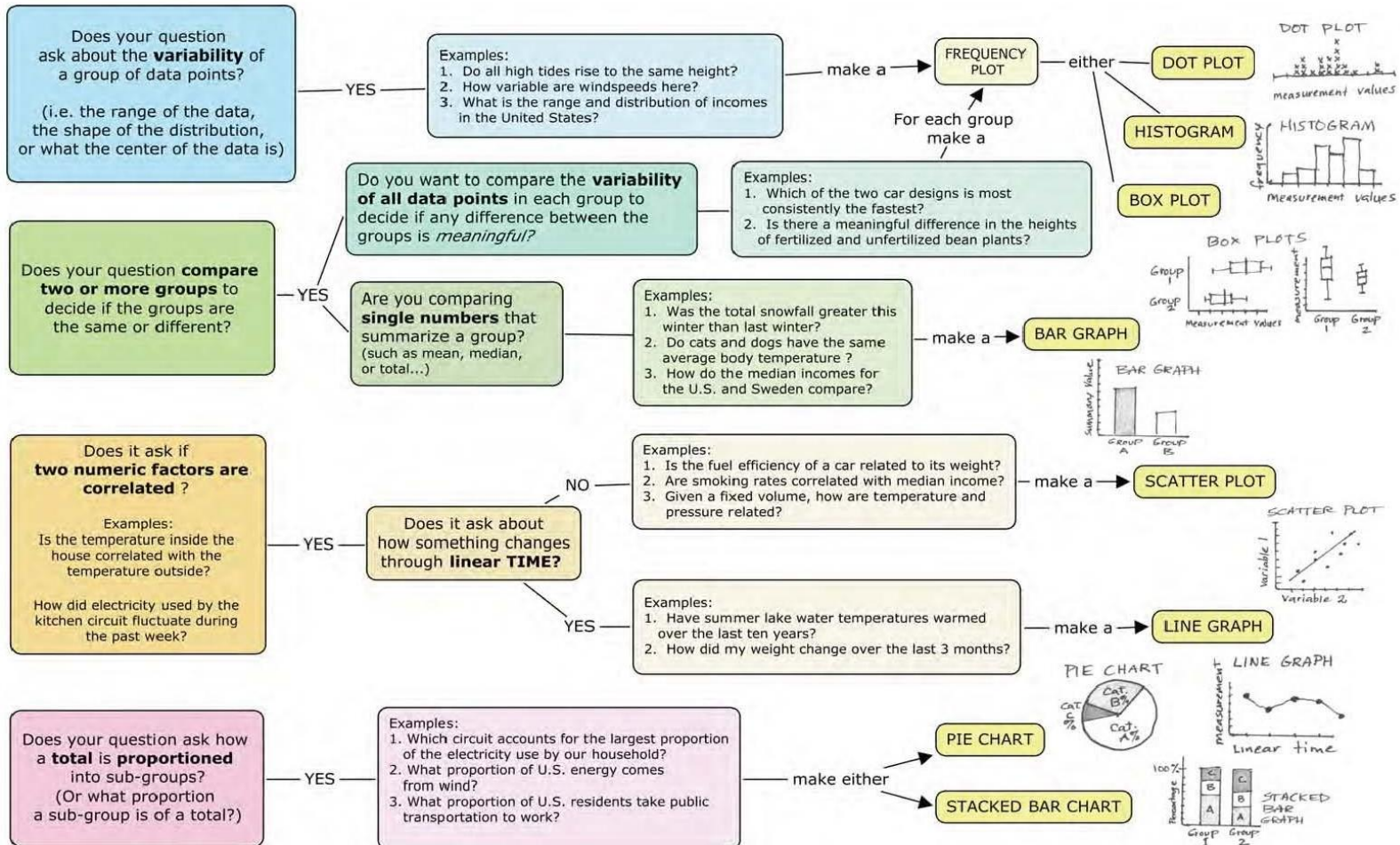
follow-up interviews it became apparent that students do not think about the type of graph that they should use to display a given set of data (Webber et al., 2014). Emphasis on the research question from which a graph is derived became the basis for the GCC. As such, this tool can be utilized as an intervention to help students in both determining what kind of information a figure is displaying and in selecting an appropriate graph to display data.



# Graph Choice Chart

What question would you like to explore?

Write your question as a complete sentence.



The Maine Data Literacy Project -- Graph Choice Chart

© 2011 The Maine Data Literacy Project (Rev. Mar 2011)

Figure 1. Graph Choice Chart. (The Maine Data Literacy Project, 2011)

Figure 1. Continued

# Graphing tips

Variability questions: Frequency plot (3 kinds)		Dot plot	Box & whisker plot	Histogram
Kind of data: <b>One categorical group</b> and <b>One numeric variable</b> (one axis)  <i>Frequency plots show how variable the group is. Describe variability by range, measure of center (mean, median, or mode), and the shape of the distribution.</i>				
Comparing groups questions:		Frequency plots	OR	Bar graph
Kind of data: <b>Two or more categorical groups</b> & <b>One numeric variable</b>  <i>Frequency plots allow you to compare how variable the groups are. Bar graphs only show a single number (ie. sum, average, percent or count) for each group.</i>		 (To compare two groups of values)		 (To compare two summary values)
Correlation questions:		Scatter plot	OR	Line graph (for time series)
Kind of data: <b>Two numeric variables</b>  <i>Both variables must be continuously numeric. Connect dots only if one variable is linear time (i.e. days, years...) Put time on the X-axis. Show correlation with a 'line of best fit'.</i>				
Proportion (percentage) questions:		Pie chart	OR	Stacked bar graph
Kind of data: Size of a subgroup as a <b>percentage</b> of the whole group (Total of sub-groups must = 100%)  <i>In pie charts and stacked bar graphs, all sub-group percentages must total 100%.</i>				

## Criteria for an informative graph:

- \_\_\_ Graph type fits the question
- \_\_\_ Axes are drawn & scaled correctly
- \_\_\_ Axes are labeled clearly, correctly
- \_\_\_ Units are given
- \_\_\_ Data are plotted accurately
- \_\_\_ Legend is present, if needed
- \_\_\_ Graph is overall neat & legible
- \_\_\_ Title and/or caption present
- \_\_\_ Trend line shown (scatter plot or line graph only)
- \_\_\_ Graph helps answer the question

(There are other kinds of questions and other kinds of graphs, and often more than one graph type is useful for a given question. Learn to graph data for these basic kinds of questions first.)

Three weeks into the course after all surveys were submitted and interviews had been conducted, the first intervention involving the GCC was implemented during lecture. The intervention involved an approximately 20 minute interactive lecture that I delivered to begin the class. Prior to the lecture, all students were given a copy of the GCC (Figure 1). The lecture engaged students by questioning how they present data and how they choose the format in which to display it. Then they used clickers to answer questions about the appropriate graph to use in a given scenario without prior instruction on how to use the GCC. I then went through the different graph options giving students the opportunity to defend their answer choice using the information on their copy of the GCC. At the end of the 20 minute intervention, students were asked the same clicker question again, after familiarizing themselves with the options using the GCC and having received direct instruction. The results were recorded. The iClicker system sent the data to an online class forum where students get participation points and the results could be viewed by instructors.

#### **3.3.4. Intervention 2: Data Summits**

The Data Summits were previously designed EES 100 assignments that involved interactive class activities. The activities were developed to engage the students with graphical material while investigating source reliability. During the semester, two different summits were run. The first Data Summit started by introducing students to different graphs related to natural resources in Maine. Students were asked to answer the question: “Is Maine pristine?” (Appendix E). For this summit, students worked in groups of about 4–8 students during class to analyze one graph out of a set of graphs gathered from websites and scientific journal articles, which was curated by the course instructors. They then completed the summit assignment as homework and brought the worksheet to the following class to engage in group discussions.

This intervention was intended to reinforce the previous intervention, giving students the opportunity to practice the skill of graph choice while targeting the other science literacy skills. One

question asked students to use the Graph Choice Chart as a guide for determining if the appropriate graph was used and to share the correct steps through the chart to get to the appropriate graph. This part of the intervention was intended to target the skill “make a graph” (Gormally et al., 2012) by engaging students in graph design. This intervention encouraged students to think critically not only about the data presented but whether or not they thought it was presented in a way that was appropriate for the type of data that were collected. Since not all representations of data support drawing accurate conclusions, especially in the case where the summary statistics and figures conflict with one another (Weissgerber et al., 2015), it is a useful skill for students to first identify if the data have been presented in an appropriate figure.

Two of the questions on the Data Summit assignment required students to determine the reliability of the source by identifying where the source originated and then evaluating it by considering possible outside (bias) influence. This was specifically designed to target the skill: conduct an effective literature search of sources and distinguish between types of sources (Gormally et al., 2012). Lastly, question six on the assignment brought in the other aspect of graphing skills: justify inferences, predictions and conclusions based on quantitative data (Gormally et al., 2012). In this question, students used graphical data to support or refute a hypothesis. This intervention was preceded by an interactive group discussion of the graphs during class where students could utilize their peers, professors, and teaching assistants to help interpret the graphs. Following the class students completed the intervention assignment on their own.

A second summit was held during the third to last week of the semester which investigated the question, “Are we in the Anthropocene?” (Appendix F). This summit, like the first, involved the use of graphs to answer the target question. Unlike the first summit, students were not supplied with graphs; rather they had to do research and discover a relevant graph that they could use to answer the question. Question one was modified to have students share a copy of the graph they selected and

question six was modified so that the hypothesis related to the question. Instead of priming the students with class discussion on the graphs, students found the graphs outside of class and brought their assignments to class to share and discuss their findings. This interaction along with feedback from the original summit assignment aided in student understanding of the targeted skills based on conclusions drawn by Lipnevich and Smith (2009) where findings showed that written feedback provided students with additional opportunities to learn and improve.

### **3.3.5. Intervention 3: Stakeholders in Research Role Play**

The last skill addressed by intervention in this study was to “distinguish between types of sources; identify bias, authority, and reliability” (Gormally et al., 2012). The focus was on student recognition that evaluation of resources by unbiased third party experts is a factor of great importance when determining the reliability or trustworthiness of a scientific paper or information source. To do this we modified a previous role play class activity that originally was designed to bring in the concept of stakeholders. The roles and the script template (Appendix G) were designed to facilitate student discussion of an environmental and human population problem from different perspectives. Specifically, students worked in small groups assigned different stakeholder roles and based on those roles, students had to identify the stance they wanted to take regarding DDT (pesticide) use in India. After convening with group members, one student from each group went to the front of the lecture hall to role play the stakeholder utilizing the group discussion in stating their stance on DDT and deliberating with the other stakeholder representatives. Then, during a Think, Pair, Share (TPS) they were asked to discuss whether any of the roles were examples of unbiased parties and then more broadly investigate what it means to be an unbiased party. This migrated from a small group discussion to a full classroom reflection.

### **3.4. Data Quality Assurance (QA), Preparation, and Statistical Analyses**

Once the pre- and post-survey results were received, the data were de-identified by an outside party associated with the distribution of the survey. The de-identified data were first quality assured by

(1) removal of any surveys completed in under 15 minutes (2) removal of any incomplete surveys where more than five questions were unanswered and (3) removal of replicate surveys. Step one in the data QA process was determined based on five volunteer surveys in which volunteers were asked to read through every question and answer, and then select a random answer from the choices. The average of these five surveys was just under 15 minutes and was therefore rounded upward. Assuming it takes approximately 15 minutes to read through the survey in order to select an answer, any survey under that time was considered to be completed for the purpose of credit and not a successful completion. This measure was used as the survey was not designed with a function to eliminate extraneous surveys. Step two was also used for elimination of extraneous surveys. Credit was given for the submission of the survey, but did not require the completion of all questions. As a result, surveys with more than five unanswered questions were considered incomplete and removed to ensure more meaningful results. Thereafter, any surveys with five or less questions unanswered, the unanswered questions were considered answered incorrectly. Step three was necessary due to the limitations of the survey software. Occasionally there was more than one survey presented for a single student. When more than one survey was present for an individual, then the most complete (i.e. one with fewer questions unanswered) was used; if both surveys were complete, the first submitted survey was used in the data set.

After the quality assessment, the data were first analyzed as a whole for the purpose of providing the EES 100 professors with the largest dataset for course evaluation. Descriptive statistics (mean, median, standard deviation, variance, minimum score, and maximum score) were calculated as a whole, as well as in subgroupings to see if there were pre-post changes on a smaller demographic scale for genders, majors, or year in school. For the purpose of this study, these data are not included as they do not provide a sample that is usable for statistical analysis, but they can be found in Appendix I.

Once the overall dataset was analyzed, matched sets were formed in order to run analyses to test for significance between the pre- and post-surveys. The overall matched data were run using a paired t-test on the difference of scores. The demographic data were compared using independent t-tests for unequal sample size and unequal variances, calculated using the difference of pre- and post-survey scores (post survey score (%) – pre-survey score). The assumption of normality for the distribution of the changes in score required for parametric analyses was met as the data exhibited a negligible right skew in the difference of the scores based on a histogram centered at 0 (i.e. no score change).

To further investigate the questions of interest (Q15, Q17, Q22, and Q28), separate analyses were run for paired nominal data using the McNemar's test (Adedokun and Burgess, 2012). The McNemar's test is non-parametric, thus the slight right skew of the observed data distribution is not problematic. The McNemar's test uses the Chi-square test statistic to determine whether proportions of response have changed across two points in time within the same population—a useful test when the data have a binary score. Since other questions on the test not specifically targeted by interventions or interviews were associated with the skills targeted, McNemar's tests were run for Q10, Q12, Q21, Q25, and Q26 as well. To determine the impact the interventions on each of these question, participation in the intervention was compared to the score changes of the matched data using a Chi-squared test. For each intervention there were two categories of students: “participated in the intervention” and “did not participate in the intervention”. These categories were compared using the subcategories associated with each question targeted by the intervention: “improved” (went from incorrect answer on pre-survey to correct answer on post-survey), “worsen” (went from correct answer on pre-survey to incorrect answer on post-survey), “no change incorrect” (incorrect on pre-survey and post-survey), and “no change correct” (correct on pre-survey and post-survey). Finally, a linear regression was run using

square root transformed data to identify if time spent on the survey (as identified in the output from open to close on the online platform) was correlated with student scores.

The interviews required a qualitative method of analysis in which participant responses were coded as either correct or incorrect. For each question in each interview, coding categories were developed to describe the type of explanation given by the student (ex. “Unbiased is reputable” or “Graph seems specific and detailed”). Responses were separated by question and the synopses were reviewed for correct and incorrect responses separately. Commonalities in each were noted and investigated further to determine why students may have answered in the way that they did. Finally, comparisons between questions that focused on the same skill (Q17 and Q22) or similar skills (Q15 and Q28) were made to see if there were any similarities or differences between student responses.



## CHAPTER 4

### RESULTS

#### 4.1. Measuring Improvement Over the Course

Of the 88 pre-surveys and 56 post-surveys that remained post-QA measures, 44 of the participants had matched data for comparison. Appendix H shows the descriptive statistics for only the matched data both overall and broken down by demographic groupings. The findings for the matched set were similar to the results of the unpaired dataset (Appendix I). The overall average scores for the pre- and post-surveys in this table are the most useful given that the demographic breakdown greatly reduces the sample size. Both the pre- and post-survey had approximately a 1:1 ratio of males to females, 1:3 ratio of EES to Non-EES majors, and 2:1 ratio of first-year students to upperclass students (Appendix H). The mean and median are relatively similar for each subset; however, the large standard deviations (roughly 20%) were explained by the bimodal distribution of pre-survey scores and the slight right skew of post-survey scores.

For students who took both the pre- and post-survey, the aggregate ( $n = 44$ ) scores did not change by the end of the course ( $p = 0.82$ ). However, when investigating sub-groups improvement was found between the pre- and post-survey of women ( $p = 0.04$ ) and the pre-post change differences were statistically significantly greater for women over men ( $p = 0.03$ ; Table 4). Men ( $n = 18$ ) earned lower overall scores on the post-survey (62%) compared to the pre-survey (66%), whereas women ( $n = 24$ ) earned higher overall scores on the post-survey (64%) compared to the pre-survey (58%). The same was true within EES ( $p = 0.04$ ) and between EES and non-EES majors ( $p = 0.05$ ; Table 5). The overall scores of non-EES majors decreased from pre-survey (63%) to post-survey (61%) whereas, EES majors increased from pre-survey (63%) to post-survey (68%; Table 5). There were no differences found within or between years in college (Table 6). All of these results are comparable to the gains found by Gormally et

al. (2012) at the classroom scale where there were significant gains of anywhere from 6-9% increases pre- to post-survey, as well as classes that showed no significant improvements.

**Table 4.** Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between men and women.

	Men	Women	Men vs Women
<b>Pre mean (%)</b>	66	58	-
<b>Post mean (%)</b>	62	64	-
<b>Change (%)</b>	-4	+6	-
<b>df</b>	17	23	11
<b>t Stat</b>	1.20	-2.15	-2.31
<b>p value</b>	0.25	0.04 <sup>s</sup>	0.04 <sup>s</sup>

TOSLS results from EES 100 course, Spring 2018, University of Maine.

<sup>s</sup> Statistically significant.

**Table 5.** Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between EES and non-EES majors.

	EES*	Non-EES**	EES vs Non-EES
<b>Pre mean</b>	61	63	-
<b>Post mean</b>	68	61	-
<b>Change</b>	+7	-2	-
<b>df</b>	11	31	29
<b>t Stat</b>	-2.31	0.57	-1.40
<b>p value</b>	0.04 <sup>s</sup>	0.57	0.17

TOSLS results from EES 100 course, Spring 2018, University of Maine.

\*Ecology and Environmental Science majors.

\*\*Majors other than EES (Non-EES).

<sup>s</sup> Statistically significant.

**Table 6.** Test of Science Literacy Skills (TOSLS) score comparison pre-post survey between first-years and upperclass students.

	First-year	Upperclass	First-year vs Upperclass
<b>Pre mean</b>	61	66	-
<b>Post mean</b>	63	63	-
<b>Change</b>	+2	-3	-
<b>df</b>	29	13	19
<b>t Stat</b>	-1.40	1.03	-1.52
<b>p value</b>	0.17	0.32	0.15

TOSLS results from EES 100 course, Spring 2018, University of Maine.

In addition, none of the focal questions examined showed significant differences between pre-survey and post-survey responses ( $p > 0.05$  for all questions; Table 7). These results were consistent both with and without the Yate's correction (a value used to reduce error). The majority of the students had no change in their response (either answered both correctly or both incorrectly); across the four targeted questions 57–84% of responses fell in this category. However, it is interesting to note that although not the target questions, improvements were seen on Q10 ( $p = 0.005$ ) and Q26 ( $p = 0.02$ ) which fall in the same skill as Q17 and Q22: 'Conduct an effective literature search of sources and distinguish between types of sources.' The last question associated with that skill, Q12, did not show significant improvement ( $p = 1.0$ ) same as Q17 and Q22. There were no other questions associated with the same skill as Q15 (make a graph), but Q21 and Q25 were aligned with the same skill as Q28 (Justify inferences, predictions, and conclusions based on quantitative data). Neither Q21 nor Q25 showed significant improvement ( $p = 0.3$ ,  $p = 0.6$  respectively).

Furthermore, participation in the Data Summits (incorporating the GCC) and role play interventions had no positive significant impact on the scores for the individual question (Table 8). In fact, students decreased in performance on question 22 ( $p = 0.03$ ; Table 8) according to participation, which is likely due to the high rates of performance overall on that question. Based on linear regression, I did not find support that the time a student spent taking the test was related to his/her score for the

pre-survey ( $F = 0.91$   $p = 0.35$ ,  $r^2 = 0.02$ ) or the post survey ( $F = 2.5,9$   $p = 0.12$ ,  $r^2 = 0.06$ ), although it should be noted that time was recorded from start to finish rather than active time.

**Table 7.** Test of Science Literacy Skills (TOSLS) question analysis McNemar results.

	<b>Q15</b>	<b>Q17</b>	<b>Q22</b>	<b>Q28</b>
<b>McNemar Chi square</b>	0.00	2.57	0.14	0.89
<b><i>p</i>-value</b>	1.00	0.11	0.71	0.35
<b>McNemar Chi square (Yate's correction (0.5))</b>	0.02	2.16	0.04	0.68
<b><i>p</i>-value (Yate's Correction)</b>	0.90	0.14	0.85	0.41

EES 100 course results, Spring 2018, University of Maine.

**Table 8.** Test of Science Literacy Skills (TOSLS) question analysis intervention participation Chi-squared results.

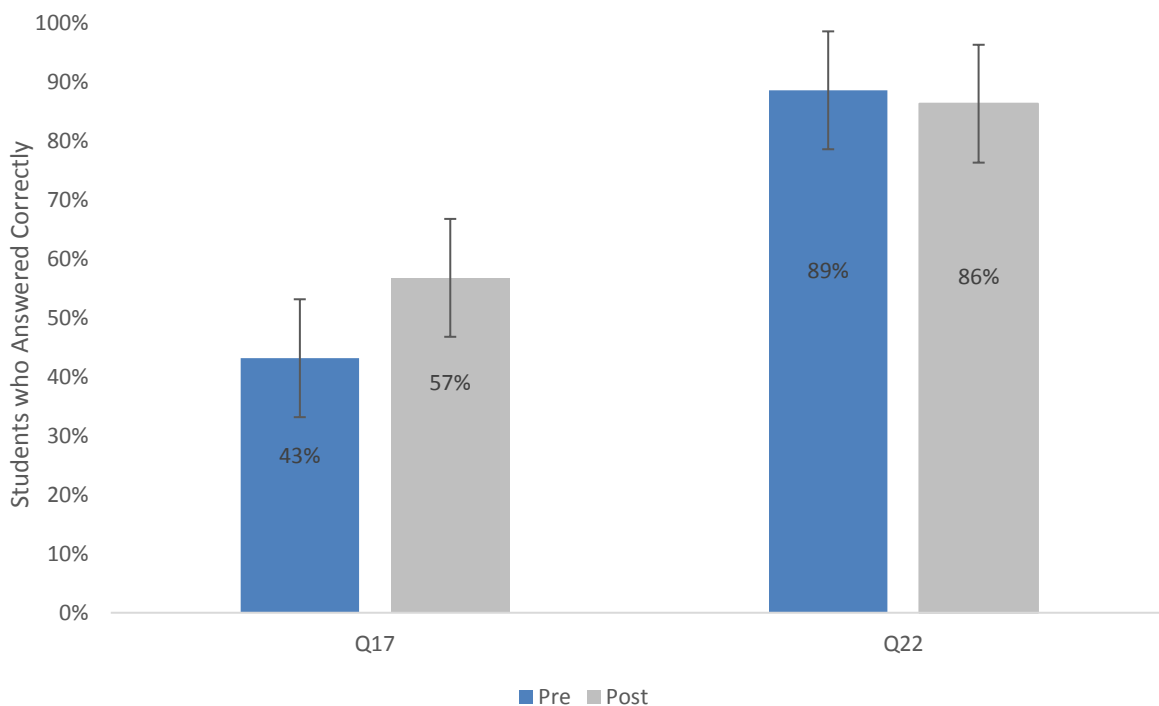
	<b>Q15</b>	<b>Q17</b>	<b>Q22</b>	<b>Q28</b>
<b>Data Summits (<i>p</i>-value)</b>	0.60	0.13	0.71	0.76
<b>Role Play (<i>p</i>-value)</b>	-	0.21	0.03 <sup>s</sup>	-

EES 100 course results, Spring 2018, University of Maine.

<sup>s</sup> Statistically significant.

## 4.2. Source Evaluation

Intervention 2 (Data Summits) and Intervention 3 (Role Play) were intended to target the skill of evaluating sources specifically Q17 and Q22 on the TOSLS survey. Although the same skill, it can be seen throughout the surveys and interviews that students were able to answer Q22 correctly more often than Q17 (Figure 2). In the initial development of TOSLS, Gormally et al. (2012) evaluated item difficulty for each question based on pilot surveys. Q22 was scored 0.72 and 0.84 on a 0–1 scale where higher values indicate an ‘easier’ question in the pre- and post-surveys respectively. This was based on a combination of the non-science major and science major courses that they evaluated. By contrast, difficulty scores for Q17 were 0.44 and 0.47 for the pre- and post-surveys respectively. Therefore, we expected variability in scores on these two questions.



**Figure 2.** Test of Science Literacy Skills (TOSLS) pre- and post-survey scores for Q17 and Q22 in EES 100, Spring 2018, University of Maine.

#### 4.2.1. Question 17

Students had no statistical gains from pre- to post- survey on Q17 ( $p = 0.11$ ; Table 7). Even though there were no gains, Q17 had the greatest positive change from incorrect on pre-survey to correct on post-survey of all the target questions. As seen in Table 9, ten students went from the incorrect answer to the correct answer with only four students having gone from the correct to the incorrect answer. Nonetheless, 15 students remained incorrect from both the pre-survey to the post-survey (Table 9).

**Table 9.** Test of Science Literacy Skills (TOSLS) Q17 McNemar’s contingency table in EES 100, Spring 2018, University of Maine.

		Post-Survey	
		# Students Correct	# Students Incorrect
Pre-Survey	# Students Correct	15	4
	# Students Incorrect	10	15

The interviews showed some trends that could explain why many students had difficulty with this question in the surveys. Of the 12 pre-interview volunteers, 58% answered correctly and of the five post-interview volunteers, 60% answered correctly. Of the participants that answered incorrectly, four (three in the pre-survey and one in the post-survey) students gave the answer “D”. However, the misconceptions that students had on Q17 (Figure 3) were overall not consistent. One participant pointed out that when collecting research they should look at the publisher because *“when growing up, we were always told in school it is about who publishes the article, you can contact them...”* Whereas another participant showed expectations that may lead to misconception such as, *“...I think they would retract anything if it was out there long enough and incorrect. I trust the publisher to do its job.”*

- 
- 17.** *The most important factor influencing you to categorize a research article as trustworthy science is:*
- a. the presence of data or graphs*
  - b. the article was evaluated by unbiased third-party experts*
  - c. the reputation of the researchers*
  - d. the publisher of the article*
- 

**Figure 3.** Test of Science Literacy Skills (TOSLS) Q17 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: ‘B’).

#### 4.2.2. Question 22

Similar to Q17, students had no statistical gains from pre- to post-survey on Q22 ( $p = 0.71$ ; Table 7). The results show a similar portion of students changed from the correct answer to incorrect (4) answer from the pre- to post-survey as changed from incorrect to correct (3; Table 10). However, 35

students remained correct in their answer to Q22 across the two surveys; this question had the greatest correct responses of all four target questions (Table 10).

**Table 10.** Test of Science Literacy Skills (TOSLS) Q22 McNemar’s contingency table in EES 100, Spring 2018, University of Maine.

		Post-Survey	
		# Students Correct	# Students Incorrect
Pre-Survey	# Students Correct	35	4
	# Students Incorrect	3	2

In both the pre- and post-interview, almost 100% of the students answered Q22 (Figure 4) correctly. In fact, the only volunteer who answered Q22 incorrectly in an interview was the individual who had not taken the survey prior to the interview. This individual also took the post-interview and at that time answered correctly. Looking at all the correct interview responses, a trend in the verbiage that students used to explain their answer arose:

*Participant 2: “...you always want an **unbiased**, well, outside opinion on [medical choices]...”*

*Participant 3: “...All the other [answer options] could have **bias**...and just a research study seems like the most effective way to find an **unbiased** outcome.”*

*Participant 7: “...a research study conducted by outside researchers would be fairly **unbiased** and would actually have decent results to share...”*

*Participant 13: “...it is an **unbiased** third party doing the research on that drug, whereas all the other things are just less trustworthy, **not necessarily unbiased**.”*

As can be seen (Figure 4), nowhere in the context of Q22 were the terms “bias” or “unbiased” used, yet 82% (pre: 10/12, post: 4/5) of the participants used this terminology to defend their answer. Only 29% (pre: 4/12, post: 1/5) of participants utilized the context of unbiased, a term found in the text of Q17 (Figure 3), in their interview responses to that question (Table 11). The use of the term “bias” or “unbiased” suggested that students understood an important component of the peer review process

which is a defining characteristic of source evaluation. However, none of the students gave a definition of bias when presenting their reasoning and therefore it is uncertain whether the students actually had an understanding of what it means to be biased or unbiased. Further inquiry in future interviews would be necessary for conclusive results.

Although students were not specific in their explanation of bias, there were other differences in student responses to Q17 and Q22 that could explain the difference in difficulty. Q17 (Figure 3) provides answer responses that are arguably all important aspects of a trustworthy article. As such, students are struggling to identify what is the most important of all the positive characteristics of a research article. Q22, on the other hand, has a set of negative buzzwords within the results. Students targeted words such as “news” or “manufacturer” as having a negative connotation when determining trustworthy aspects of research. Students were more apt to remove responses that they associated with negative aspects of research than they were to eliminate responses that contained terms that had positive associations with trustworthy science.

---

*22. Your doctor prescribed you a drug that is brand new. The drug has some significant side effects, so you do some research to determine the effectiveness of the new drug compared to similar drugs on the market. Which of the following sources would provide the most accurate information?*

- a. the drug manufacturer's pamphlet/website*
  - b. a special feature about the drug on the nightly news*
  - c. a research study conducted by outside researchers*
  - d. information from a trusted friend who has been taking the drug for six months*
- 

**Figure 4.** Test of Science Literacy Skills (TOSLS) Q22 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'C').



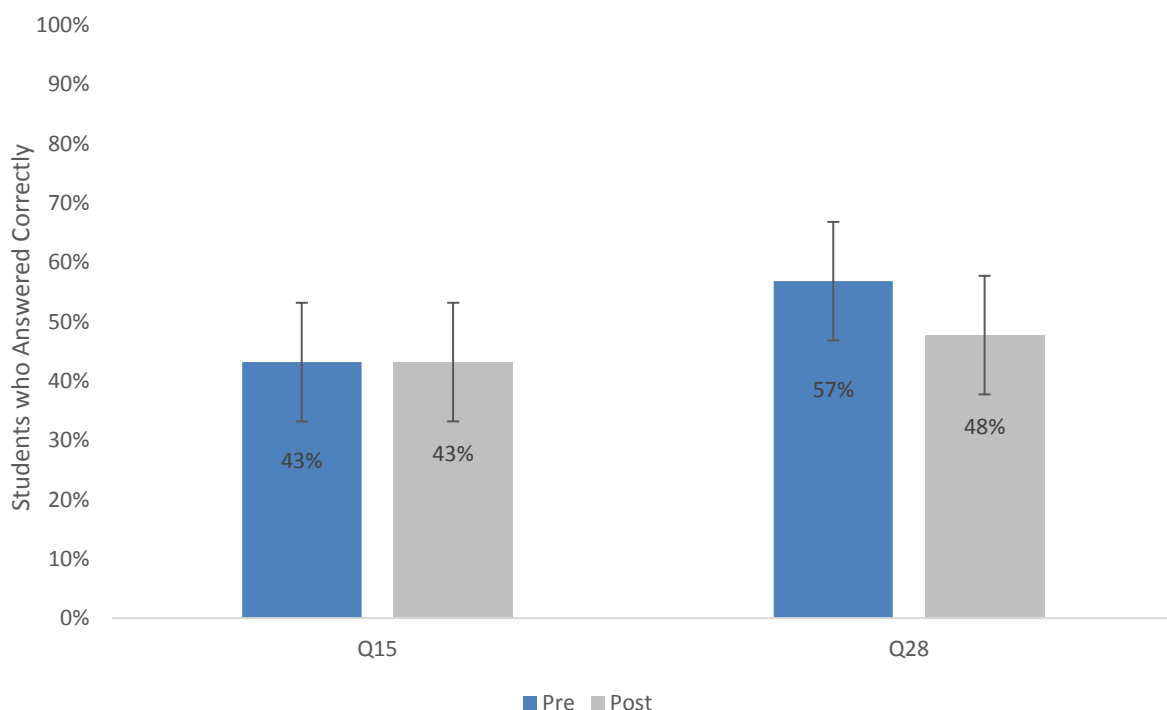
**Table 11.** Interviewee use of the term “bias” or “unbiased” in Test of Science Literacy Skills (TOSLS) Q17 and Q22 for EES 100, Spring 2018, University of Maine.

Interview	1	2	3	4*	5	6	7	8	9	10	11	12	13	14	15	16	17
Q17	No	Yes	Yes	No	No	No	Yes	No	Yes	No	No	No	No	No	Yes	No	No
Q22	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes

\*This participant did not take the pre-survey prior to the interview

### 4.3 Graphical Representation and Interpretation

Intervention 1 (GCC) and Intervention 2 (Data Summits) were intended to target the skills of making a graph (Q15) and justifying inferences, predictions, and conclusions based on quantitative data (Q28). Both Q15 and Q28 involve evaluation of graphs provided in the survey questions. Although not the same skill, both rely heavily on understanding and interpreting the information provided by graphs. When measuring the difficulty of these two skills, Gormally et al. (2012) both fell at or below 0.5 on the difficulty scale in both pre- and post-surveys indicating that the two are among the more difficult questions. As such, lower overall scores on these questions seen in Figure 5 were anticipated, as was the case with Q17.



**Figure 5.** Test of Science Literacy Skills (TOSLS) pre- and post-survey scores for Q15 and Q28 in EES 100, Spring 2018, University of Maine.

#### 4.3.1. Question 15

Changes in pre/post results on the surveys were not statistically significant for Q15 ( $p = 1.00$ ; Table 7). As can be seen in Table 12, the same number of students went from the incorrect answer on the pre-survey to the correct answer on the post survey as went from correct to incorrect. This inconsistency suggests that students were having difficulties understanding the question enough to stay confident in their answers. This is also supported by the overall average (43%, Figure 5) which shows that fewer than half of the students were correct on both surveys.

**Table 12.** Test of Science Literacy Skills (TOSLS) Q15 McNemar contingency table in EES 100, Spring 2018, University of Maine.

		Post-Survey	
		# Students Correct	# Students Incorrect
Pre-Survey	# Students Correct	11	8
	# Students Incorrect	8	17

Only 17% of interviewees answered Q15 correctly on the pre-interview and none of the post-interview participants answered correctly. Of the participants who answered Q15 (Figure 6) incorrectly, 80% chose graph C. According to participant responses, many selected graph C because they felt more information was provided— given the individual points— than in the other graphs.

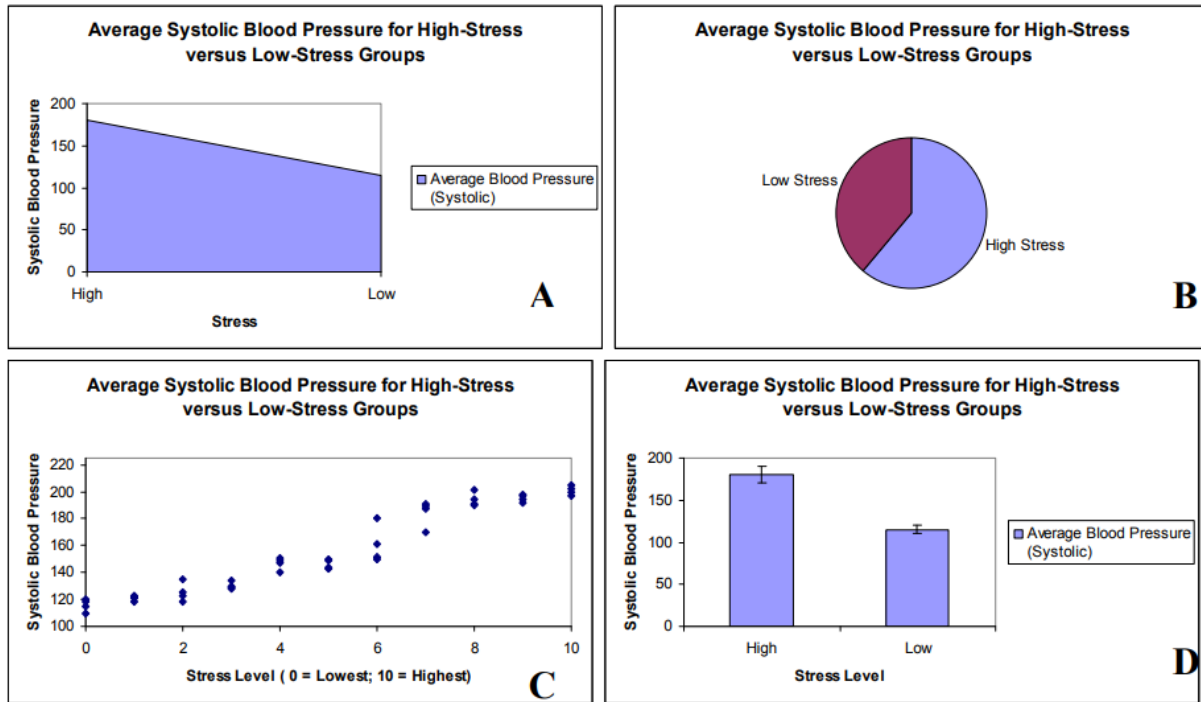
**Participant 3:** *“It shows a range of stress levels so it is just more specific...”*

**Participant 5:** *“...[the points] are not all connected, [the graph] shows individual points...”*

**Participant 6:** *“...there was more data represented in [graph C], like different stress levels versus blood pressure, whereas in this one it is just high and low...”*

Although they vary in description, the overarching theme from respondents for choosing the incorrect answer was that more information was given in graph C since the individual data points were shown rather than a summary of the data from which individual responses could not be discerned.

15. Researchers found that chronically stressed individuals have significantly higher blood pressure compared to individuals with little stress. Which graph would be most appropriate for displaying the mean (average) blood pressure scores for high-stress and low-stress groups of people?



**Figure 6.** Test of Science Literacy Skills (TOSLS) Q15 used in EES 100 Spring 2018 interviews, University of Maine (Answer: 'D').

#### 4.3.2. Question 28

The difference in student performance from pre-survey to post-survey was also not statistically significant for Q28 ( $p = 0.35$ ; Table 7). Eleven students went from the correct answer in the pre-survey to the incorrect answer in the post survey (Table 13). Only seven students improved pre-survey to post-survey. This inconsistency in responding correctly again identifies a student difficulty with this question which is also shown by the approximately 10% decrease in score for Q28 from the pre-survey to post-survey (Table 13).

**Table 13.** Test of Science Literacy Skills (TOSLS) Q28 McNemar’s contingency table in EES 100, Spring 2018, University of Maine.

		Post-Survey	
		# Students Correct	# Students Incorrect
Pre-Survey	# Students Correct	14	11
	# Students Incorrect	7	12

In response to Q28 (Figure 7) in the interviews, none of the participants selected answer choice “A”; some stated that the option was irrelevant to the question while others simply did not mention it. Similar to the survey responses, the post-interview average (20%) was less than the pre-interview average (58%). Across all the interviews, Q28 was answered correctly (“B”) by 47% of the interviewees, however, 41% of interviewees were stumped by answer “D”. Students that chose this answer did not seem to notice that the direction of the trend was opposite the narrative description. Rather, those who answered incorrectly seemed to be focused on the hypothesis that supplemented the trend and whether or not that hypothesis seemed correct. Students who answered “D” supported it with statements such as:

**Participant 4:** *“I picked ‘D’ because it kind of does make sense that there are a lot less predators at the higher elevation as you go up. A lot of things can’t survive.”*

**Participant 8:** *“You can see that as the elevation goes up you have smaller shrimp population. Less shrimp so you know that at a higher elevation they probably have less predators.”*

**Participant 12:** *“I would go with ‘D’ because I do think they would have fewer predators at higher elevations. I can’t imagine many things using so much energy to get so high up to eat a little thing.”*

---

**Background for question 28:** Researchers interested in the relation between River Shrimp (*Macrobrachium*) abundance and pool site elevation, presented the data in the graph below. Interestingly, the researchers also noted that water pools tended to be shallower at higher elevations.

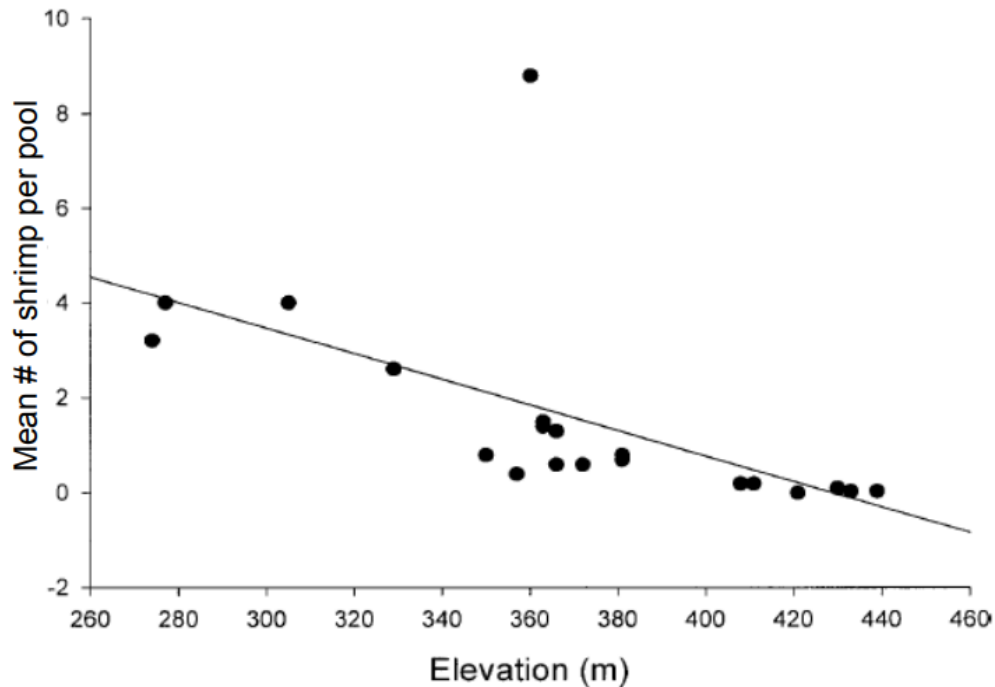


FIG. 3. Relationship between total abundance of *Macrobrachium* (1988–2002) and elevation in Quebrada Prieta.

**28.** Which of the following is a plausible hypothesis to explain the results presented in the graph?

- a. There are more water pools at elevations above 340 meters because it rains more frequently in higher elevations.
- b. River shrimp are more abundant in lower elevations because pools at these sites tend to be deeper.
- c. This graph cannot be interpreted due to an outlying data point.
- d. As elevation increases, shrimp abundance increases because they have fewer predators at higher elevations.

---

**Figure 7.** Test of Science Literacy Skills (TOSLS) Q28 used in EES 100 Spring 2018 interviews, University of Maine (correct answer: 'B').

## CHAPTER 5

### DISCUSSION

#### 5.1. Conclusions of the Study

The interventions in this study were designed as the first stepping stones in a process from which we can ultimately derive educational resources to implement throughout science curricula. Although my hypothesis that these interventions would show gains in TOSLS scores – in particular with the questions of interest – was not supported overall or in the majority of demographic groups, much was learned about how students are thinking around the science literacy topics of source and graphical evaluations. This serves as a starting point for intervention modifications. Lack of participation in the post-survey interviews meant that there was not a sufficiently large sample of the population (5 interviews out of 148 students) to utilize the information obtained on questions designed to provide students' feedback on the interventions.

Rather than seeing the impact of the interventions on student proficiency based on TOSLS, I found from the interviews that students struggled with some similar types of misconception both before and after the interventions. When faced with determining the reliability of a source, students were aware that there are important aspects to look for such as a reliable publisher and researcher as well as supportive data and figures. However, the students could not always pinpoint the factor of an “unbiased third party expert” as one of the strongest components of trustworthy science, despite having participated in an in-class role play activity that used that language. As one study involving six disciplines in 50 colleges with responses from 278 educators highlighted there are discrepancies in teaching about information literacy (source evaluation). It was found that educators differed in their perceptions of the importance and relevance of information literacy in their courses (Saunders, 2012). As such, students across classes and institutions are not being provided with the same information. With consistent

methods of assessment and targeted intervention, these discrepancies can be mitigated and concepts can be reinforced throughout primary and secondary education.

Interviews also uncovered the difficulties that students had with the science literacy skills involving quantitative data and graphing. Interviews showed that students got caught up in the quantity of data provided and they were therefore unable to reconcile what graph was most appropriate for the question at hand. A similar trend was shown in a study investigating quantitative literacy where science-major students in a university introductory biology course were unable to select the appropriate graph to utilize when asked to graph a set of data (Bray Speth et al., 2010). Identifying the research question is the first step in determining graph configuration (Webber et al., 2014). Without this step, any type of graph could potentially be used to display the data. Students must be aware of the intent of the graph in order to produce the appropriate type of graph. Not only did students struggle with the design of graphs but they also had difficulty with what the graph portrayed. Interviews revealed that justifying inferences was difficult for students as the inferences provided relied on students' own prior knowledge rather than the data presented to them. Again, this source of difficulty with graphs is well documented in the literature, even recently (Pérez-Echeverría, Postigo, & Marín, 2018). The difficulties in these skills in particular are important to note as it is necessary as a scientifically literate citizen to be able to read a graph, identify the trend, and determine if the drawn conclusions are appropriate. If we know what skills students struggle with the most, such as these specific concerns, we can design studies that will uncover students' common misconceptions or discover where they get distracted from the important features of a problem or graph. With that knowledge, we can then design better interventions to target these skills.

The interventions in this study were not able to change student thinking on these topics that have been shown to be difficult. Instead these findings suggest that greater intervention is necessary in order to see improvement in these areas (source evaluation and graph evaluation and interpretation) via instruction aimed at students' specific misunderstandings. Other factors could have driven these

results – the methods of delivery, student familiarity (or lack thereof) with the material, time spent with each skill, or difficulty with understanding, reading, or interpreting the questions. Nonetheless, by focusing on the extent to which students struggle with these science literacy skills, we can more efficiently target the intervention and eventually see improvement on TOSLS in these skill areas. Continuing to investigate the specific difficulties students have within each skill will allow for more directed educational instruction so that classroom time can be used more effectively to help students overcome previous conceptions. In the long run, this educational experience will allow for the formation of more scientifically literate citizens.

Scientific research is held to a high standard because of the rigor of the methodology, evaluation, and peer review processes that drive the many fields of research. Science literacy is at the core of this and it is an unspoken assumption that scientists inherently have these skills. Yet, when looking at the results of this study, in a course in which approximately two-thirds of students are science-related majors, the results were disappointing with an average for the course below 65% for both matched (Appendix H) and unmatched (Appendix I) data sets, indicating relatively low fluency in science literacy. At some point these skills need to be imparted upon these students who will enter their scientific fields of study with the expectation that they already have these skills. This raises two concerns: first, the prior concern with the inability to make sound scientific decisions; and second, the concern that science students do not have command of these skills early in their academic training, and thus rely on college and post-graduate experience to develop these core science skills, perhaps inadequately. This has a compounding effect when citizens do not have the ability to identify a well-constructed study from a poorly designed one (Scheufele, 2013). Therefore, reform in science literacy education starting prior to college is vital to uphold scientific rigor and ensure students have these skills internalized prior to entering their fields of study.



## 5.2. Limitations and Modifications

In order to continue progress toward a toolbox of interventions that can be used as educational resources for improving student science literacy proficiency, the limitations of this study must be noted. Results of this study were greatly limited by the participation of students on the surveys as well as in the interviews. Although the surveys were available for course participation credit to all 148 students (measured at the beginning of the semester), only 111 students submitted the pre-survey and only 90 submitted the post-survey. These numbers were further restricted by the QA methods which were deployed independently of the TOSLS instruction. Participation in surveys was even less as these opportunities were voluntary and supplied an incentive that may have had an influence on certain demographics that may not have been representative of the class population.

Of the 148 students, 12 participated in the pre-interview whereas only 5 participated in the post-interview. Majority trends in interview results suggested class-scale difficulties; however, the assumption remains that the samples were representative. Requiring interviewees to participate in both the pre-survey and post-survey was difficult given that commitments from beginning of the semester may not be upheld at the end of the semester. As was apparent in the results, procuring volunteers became more difficult at the end of the semester.

Limitations with the interventions also included time, data collection, and limited pilot data for determination of target skills. Due to time constraints of the semester, only a certain amount of time could be allotted to provide the students with opportunities to develop the target science literacy skills. Given this, some interventions occurred outside of class where students may have misinterpreted what they were supposed to do for an assigned task and as such they may not have worked towards improvement of the targeted science literacy skills. As mentioned in Falk et al. (2016), opportunities where information is presented informally can lead to misinterpretation. As such, I would recommend that future interventions be carried out in the classroom to ensure that students are presented with the

information as intended. However, I would suggest maintaining the interactive nature of the activities as were used throughout the interventions utilizing peer interaction and discussion. Volunteers for the post interviews were quick to recall the activities as they were much more memorable to them as they played a role in the education of themselves as well as their classmates. With the currently designed interventions, data collection should be deployed within the interventions as was attempted by using the iClicker system and grading assignments; however, these should be a comparable measure to relate to the TOSLS skills being targeted. Further, data could be collected from students at the end of the post-survey inquiring about the usefulness of the interventions to use as broad scale feedback rather than only the five sets of data collected in the post-interview. Student feedback throughout would be invaluable to the construction of stronger interventions.

Finally, targeting the appropriate skills for the course in question and the difficulties students have with those skills is key for building upon student understanding. In this study, information was used from a pilot year rather than the current semester. Although there were obvious struggles in the selected skills in 2018 as well, not enough was determined about the particulars of those struggles in 2017 that could lead to impactful interventions. Targeting the skills, finding the gaps in student understanding within the skill, and then providing educators with tools to fill those gaps will be the foundation for making scientifically literate individuals.

### **5.3. Implications**

Studies have shown the need for scientifically literate citizens regardless of nationality, race, ethnicity, gender, social class, or age (Allum, Besley, Gomez, & Brunton-Smith, 2018; Msafiri, 2018). Yet, there are few resources available describing methods in which to bring about these scientifically literate citizens. As Feinstein (2010) points out, there has been little impact of science education on use and interpretation of science in everyday life. So not only is there a lack of resources from which to educate people in science literacy, but those that are in practice provide little aid in the practical use of science

literacy. The interventions designed in this study were just the first step in the iterative design process to provide educators with resources to aid students in developing science literacy skills that can be utilized in daily life. The intent was to fill the need for educational materials focused in science literacy for classroom education. The need for these resources in education have been discussed by many as they recognize the need for scientifically literate citizens (Matschullat, 2015; Millar, 2006; Gormally et al., 2012; Reeves and Honig, 2015).

As instructed by Gormally et al. (2012), TOSLS was used to determine gaps in students' proficiency in certain science literacy skills as compared to the intent of the science literacy instruction, in this case the interventions. This helped determine whether or not there was a quantifiable impact of the intervention on students' proficiency in the targeted skills. By having a measurable outcome – a TOSLS score – we have provided a standard by which other interventions can be compared. As previously mentioned, TOSLS is a multidimensional tool that can be used in different post-secondary educational settings, from small colleges to large general education courses at universities. This tool also targets science literacy in a framework that is relevant to daily life and thus should improve upon the lack of relevancy that Feinstein (2010) discovered in the education of science literacy. The more a tool such as TOSLS can be used, the more data will be available to evaluate what methods of education will demonstrate an increase in student proficiency in science literacy.

## REFERENCES

- Adedokun, O. A., & Burgess, W. D. (2012). Analysis of Paired Dichotomous Data: A Gentle Introduction to the McNemar Test in SPSS. *Journal of Multidisciplinary Evaluation*, 8(17), 125-131.
- Allum, N., Besley, J., Gomez, L., & Brunton-Smith, I. (2018). Disparities in Science Literacy. *Science*, 360(6391), 861-862. <https://doi.org/10.1126/science.aar8480>
- Benjamin, T. E., Marks, B., Demetrikopoulos, M. K., Rose, J., Pollard, E., Thomas, A., & Muldrow, L. L. (2017). Development and Validation of Scientific Literacy Scale for College Preparedness in STEM with Freshmen from Diverse Institutions. *International Journal of Science and Mathematics Education*, 15(4), 607–623. <https://doi.org/10.1007/s10763-015-9710-x>
- Blank, J. M., McGaughey, K. J., Keeling, E. L., Thorp, K. L., Shannon, C. C., & Scaramozzino, J. M. (2016). A Novel Assessment Tool for Quantitative Evaluation of Science Literature Search Performance: Application to First-Year and Senior Undergraduate Biology Majors. *College & Research Libraries*, 682-702. <https://doi.org/10.5860/crl.77.6.682>
- Bray Speth, E., Momsen J. L., Moyerbrailean, G. A., Ebert-May, D., Long, T. M., Wyse, S., & Linton, D. (2010). 1, 2, 3, 4: Infusing Quantitative Literacy into Introductory Biology. *CBE Life Science Education*, 9, 323–332. <https://doi.org/10.1187/cbe.10-03-0033>
- Dasgupta, A. P., Anderson, T. R., & Pelaez, N. (2014). Development and Validation of a Rubric for Diagnosing Students' Experimental Design Knowledge and Difficulties. *CBE Life Sciences Education*, 13(2), 265–284. <https://doi.org/10.1187/cbe.13-09-0192>
- Deane, T., Nomme, K., Jeffery, E., Pollock, C., & Birol, G. (2016). Development of the Statistical Reasoning in Biology Concept Inventory (SRBCI). *CBE Life Sciences Education*, 15(1), 1–13. <https://doi.org/10.1187/cbe.15-06-0131>
- Deboer, G. E. (2000). Scientific Literacy: Another Look at Its Historical and Contemporary Meanings and Its Relationship to Science Education Reform. *Journal of Research In Science Teaching*, 37(6), 582–601. [https://doi-org.prxy4.ursus.maine.edu/10.1002/1098-2736\(200008\)37:6<582::AID-TEA5>3.0.CO;2-L](https://doi-org.prxy4.ursus.maine.edu/10.1002/1098-2736(200008)37:6<582::AID-TEA5>3.0.CO;2-L)
- Eskinasi, M, & Fokkema, E. (2006). Lessons Learned from Unsuccessful Modelling Interventions. *Systems Research and Behavioral Science*, 23, 483-492. <https://doi.org/10.1002/sres.774>
- Falk, J. H., Dierking, L. D., Swanger, L. P., Staus, N., Back, M., Barriault, C., ... Verheyden, P. (2016). Correlating Science Center Use With Adult Science Literacy: An International, Cross-Institutional Study. *Science Education*, 100(5), 849-876 . <https://doi.org/10.1002/sce.21225>
- Feinstein, N. (2010). Salvaging Science Literacy. *Science Education*, 95, 168-195. <https://doi.org/10.1002/sce.20414>

- Gormally, C., Brickman, P., Hallar, B., & Armstrong, N. (2009). Effects of Inquiry-based Learning on Students' Science Literacy Skills and Confidence. *International Journal for the Scholarship of Teaching and Learning*, 3(2). <https://doi.org/10.20429/ijstol.2009.030216>
- Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a Test of Scientific Literacy Skills (TOSLS): Measuring Undergraduates' Evaluation of Scientific Information and Arguments. *CBE Life Sciences Education*, 11(4), 364–377. <https://doi.org/10.1187/cbe.12-03-0026>
- Kiuhara, S., & Witzel, B. (2014). Math Literacy Strategies for Students With Learning Difficulties. *Childhood Education*, 90(3), 234–238. <https://doi.org/10.1080/00094056.2014.912067>
- Koltay, T. (2016). Data Governance, Data Literacy and The Management of Data Quality. *International Federation of Library Associations and Institutions*, 42(4), 303-312. <https://doi.org/10.1177/0340035216672238>
- Lawson, A. E. (1978). The Development and Validation of a Classroom Test of Formal Reasoning. *Journal of Research in Science Teaching*, 15(1), 11–24. <https://doi.org.prxy4.ursus.maine.edu/10.1002/tea.3660150103>
- Lederman, N.G. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching*, 39(6), 497-521. <https://doi.org/10.1002/tea.10034>
- Liang, L. L., Chen, S., Chen, X., Kaya, O.N., Adams, A.D., Macklin, M., & Ebenezer, J. (2008). Assessing Preservice Elementary Teachers' Views on the Nature of Scientific Knowledge: A Dual-response Instrument. *Asia-Pacific Forum on Science Learning and Teaching*, 9(1), 1-20.
- Lipnevich, A. A., & Smith, J. K. (2009). Effects of Differential Feedback on Students' Examination Performance. *Journal Of Experimental Psychology: Applied*, 15(4), 319-333. <https://doi.org/10.1037/a0017841>
- Pérez-Echeverría, M. P., Postigo, Y., & Marín C. (2018). Understanding of Graphs in Social Science Undergraduate Students: Selection and Interpretation of Graphs. *Irish Educational Studies*, 37(1), 89-111. <https://doi.org/10.1080/03323315.2018.1440248>
- Majima, Y. (2015). Belief in Pseudoscience, Cognitive Style and Science Literacy. *Applied Cognitive Psychology*, 29, 552-559. <https://doi.org/10.1002/acp.3136>
- Matschullat, J. (2015). Global Resource Science Literacy: a leading program. *Environmental Earth Sciences*, 73(12), 8709-8712. <https://doi.org/10.1007/s12665-014-3967-1>
- Millar, R. (2006). Twenty First Century Science: Insights from the Design and Implementation of a Scientific Literacy Approach in School Science. *International Journal of Science Education*, 28(13), 1499–1521. <https://doi.org/10.1080/09500690600718344>
- Morgan, R. E., Bertera, R. L., & Reid, L. A. (2007). An Intergenerational Approach to Informal Science Learning and Relationship Building. *Journal of Intergenerational Relationships*, 5(3), 27-43. [https://doi.org/10.1300/J194v05n03\\_03](https://doi.org/10.1300/J194v05n03_03)

- Msafiri. (2018) Urban Barbarians and Science Literacy. *Information Development*, 34(3), 315.  
<https://doi.org/10.1177/0266666918770384>
- Reeves, T.D. & Honig, S.H. (2015). A Classroom Data Literacy Intervention for Pre-service Teachers. *Teaching and Teacher Education*, 50, 90-101. <https://doi.org/10.1016/j.tate.2015.05.007>
- Scheufele, D.A. (2013). Communicating Science in Social Settings. *Proceedings of the National Academy of Sciences*, 110 (Supplement 3), 14040-14047.
- Schielke, K. (2013). Building Science Literacy by reading Science News. *Science Scope*, 37(3), 20-25.
- Saunders, L. (2012). Faculty Perspectives on Information Literacy as a Student Learning Outcome. *Journal of Academic Librarianship*, 38(4), 226-236. <https://doi.org/10.1016/j.acalib.2012.06.001>
- The University of Maine. (2016). 2016-2017 Undergraduate Catalog. Retrieved from  
[http://catalog.umaine.edu/preview\\_course\\_nopop.php?catoid=67&coid=193186](http://catalog.umaine.edu/preview_course_nopop.php?catoid=67&coid=193186)
- Valdecasas, A. G., & Correias, A. M. (2010). Science Literacy and Natural History Museums. *Journal of Biosciences*, 35, 507-514. <https://doi.org/10.1007/s12038-010-0057-8>
- Webber, H., Nelson, S. J., Weatherbee, R., Zoellick, B., & Schauffler, M. (2014). The Graph Choice Chart: A Tool to help Students Turn Data into Evidence. *The Science Teacher*, 81(8), 37-43.
- Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PLoS Biology*, 13(4), 1–10.  
<https://doi.org/10.1371/journal.pbio.1002128>

## APPENDIX A: TEST OF SCIENCE LITERACY SKILLS SURVEY (Gormally et al., 2012)

**Directions:** There are 28 multiple-choice questions. You will have about 35 minutes to work on the questions. Be sure to answer as many of the questions as you can in the time allotted. You will receive attendance points for completing the entire assignment today. Your grade will depend on completeness and thoroughness, not on correct answers. But, try your best, your honest answers will help us better prepare the materials for the remainder of the semester.

**Mark your answers on the scantron sheet.**

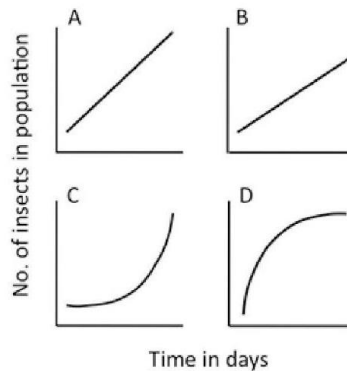
**Bubble in your #ID on your scantron.**

Do NOT use a calculator. Thank you for your participation in this project!

1. Which of the following is a valid scientific argument?
  - a. Measurements of sea level on the Gulf Coast taken this year are lower than normal; the average monthly measurements were almost 0.1 cm lower than normal in some areas. These facts prove that sea level rise is not a problem.
  - b. A strain of mice was genetically engineered to lack a certain gene, and the mice were unable to reproduce. Introduction of the gene back into the mutant mice restored their ability to reproduce. These facts indicate that the gene is essential for mouse reproduction.
  - c. A poll revealed that 34% of Americans believe that dinosaurs and early humans co-existed because fossil footprints of each species were found in the same location. This widespread belief is appropriate evidence to support the claim that humans did not evolve from ape ancestors.
  - d. This winter, the northeastern US received record amounts of snowfall, and the average monthly temperatures were more than 2°F lower than normal in some areas. These facts indicate that climate change is occurring.

2. While growing vegetables in your backyard, you noticed a particular kind of insect eating your plants. You took a rough count (see data below) of the insect population over time. Which graph shows the best representation of your data?

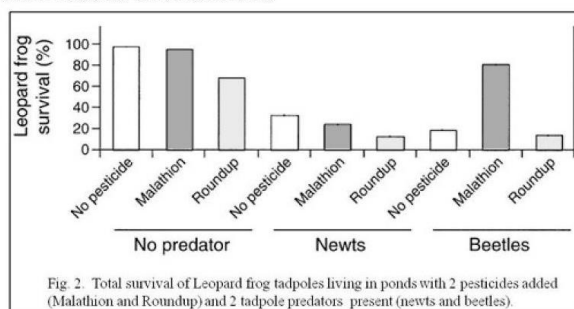
Time (days)	Insect Population (number)
2	7
4	16
8	60
10	123



3. A study about life expectancy was conducted using a random sample of 1,000 participants from the United States. In this sample, the average life expectancy was 80.1 years for females and 74.9 years for males. What is one way that you can increase your certainty that women truly live longer than men in the United States' general population?
  - a. Subtract the average male life expectancy from the average female expectancy. If the value is positive, females live longer.
  - b. Conduct a statistical analysis to determine if females live significantly longer than males.
  - c. Graph the mean (average) life expectancy values of females and males and visually analyze the data.
  - d. There is no way to increase your certainty that there is a difference between sexes.

4. Which of the following research studies is **least likely** to contain a confounding factor (variable that provides an alternative explanation for results) in its design?
  - a. Researchers randomly assign participants to experimental and control groups. Females make up 35% of the experimental group and 75% of the control group.
  - b. To explore trends in the spiritual/religious beliefs of students attending U.S. universities, researchers survey a random selection of 500 freshmen at a small private university in the South.
  - c. To evaluate the effect of a new diet program, researchers compare weight loss between participants randomly assigned to treatment (diet) and control (no diet) groups, while controlling for average daily exercise and pre-diet weight.
  - d. Researchers tested the effectiveness of a new tree fertilizer on 10,000 saplings. Saplings in the control group (no fertilizer) were tested in the fall, whereas the treatment group (fertilizer) were tested the following spring.
5. Which of the following actions is a valid scientific course of action?
  - a. A government agency relies heavily on two industry-funded studies in declaring a chemical found in plastics safe for humans, while ignoring studies linking the chemical with adverse health effects.
  - b. Journalists give equal credibility to both sides of a scientific story, even though one side has been disproven by many experiments.
  - c. A government agency decides to alter public health messages about breast-feeding in response to pressure from a council of businesses involved in manufacturing infant formula.
  - d. Several research studies have found a new drug to be effective for treating the symptoms of autism; however, a government agency refuses to approve the drug until long term effects are known.

**Background for question 6:** The following graph appeared in a scientific article<sup>1</sup> about the effects of pesticides on tadpoles in their natural environment.

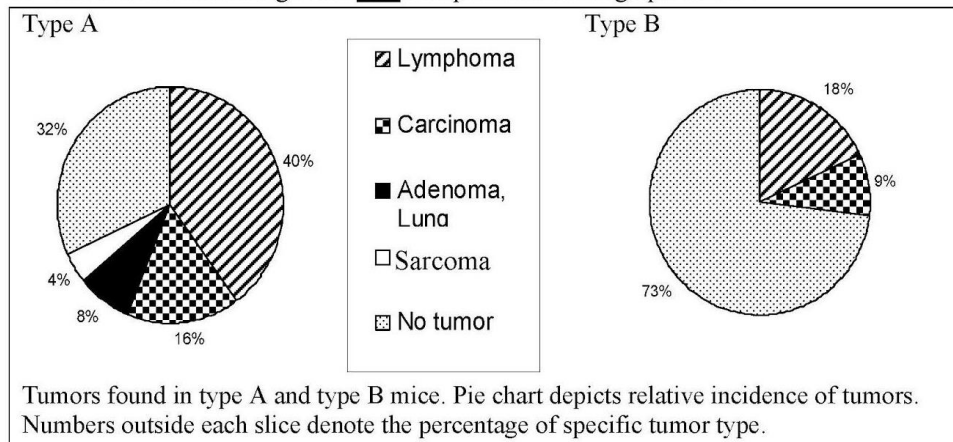


6. When beetles were introduced as predators to the Leopard frog tadpoles, and the pesticide Malathion was added, the results were unusual. Which of the following is a plausible hypothesis to explain these results?
  - a. The Malathion killed the tadpoles, causing the beetles to be hungrier and eat more tadpoles.
  - b. The Malathion killed the tadpoles, so the beetles had more food and their population increased.
  - c. The Malathion killed the beetles, causing fewer tadpoles to be eaten.
  - d. The Malathion killed the beetles, causing the tadpole population to prey on each other.

<sup>1</sup> Modified from Relyea, R.A., N.M. Schoeppner, J.T. Hoverman. 2005. Pesticides and amphibians: the importance of community context. *Ecological Applications* 15: 1125-1134



7. Which of the following is the **best** interpretation of the graph below<sup>2</sup>?



- a. Type “A” mice with Lymphoma were more common than type “A” mice with no tumors.
  - b. Type “B” mice were more likely to have tumors than type “A” mice.
  - c. Lymphoma was equally common among type “A” and type “B” mice.
  - d. Carcinoma was less common than Lymphoma only in type “B” mice.
8. Creators of the Shake Weight, a moving dumbbell, claim that their product can produce “incredible strength!” Which of the additional information below would provide the **strongest evidence** supporting the effectiveness of the Shake Weight for increasing muscle strength?
- a. Survey data indicates that on average, users of the Shake Weight report working out with the product 6 days per week, whereas users of standard dumbbells report working out 3 days per week.
  - b. Compared to a resting state, users of the Shake Weight had a 300% increase in blood flow to their muscles when using the product.
  - c. Survey data indicates that users of the Shake Weight reported significantly greater muscle tone compared to users of standard dumbbells.
  - d. Compared to users of standard dumbbells, users of the Shake Weight were able to lift weights that were significantly heavier at the end of an 8-week trial.
9. Which of the following is **not** an example of an appropriate use of science?
- a. A group of scientists who were asked to review grant proposals based their funding recommendations on the researcher’s experience, project plans, and preliminary data from the research proposals submitted.
  - b. Scientists are selected to help conduct a government-sponsored research study on global climate change based on their political beliefs.
  - c. The Fish & Wildlife Service reviews its list of protected and endangered species in response to new research findings.
  - d. The Senate stops funding a widely used sex-education program after studies show limited effectiveness of the program.

<sup>2</sup> Modified from Wang, Y., S. Klumpp, H.M. Amin, H. Liang, J. Li, Z. Estrov, P. Zweidler-McKay, S.J.Brandt, A. Agulnick, L. Nagarajan. 2010. SSBP2 is an *in vivo* tumor suppressor and regulator of LDB1 stability. *Oncogene* 29: 3044-3053.

**Background for question 10:** Your interest is piqued by a story about human pheromones on the news. A Google search leads you to the following website:

**EROS FOUNDATION**

**Special Sale**  
Pheromone 10.13™ increase romance in your life; 1.6 oz. bottle normally \$98.50. (25% off for first time customers.) [Order Now](#)

EROS HOME EROS SCIENCE PHEROMONE DISCOVERY BOOKS AND PRODUCTS MEDIA ARTICLES CONTACT US VIDEO LINKS

**Shortcuts**  
Click here To Order From Eros  
Privacy Protection  
Share

**Explore the Site**  
Eros Home  
Top Stories  
Dr. Baxter's Articles  
Discoveries  
Baxter in the Scientific Community  
Other Health Research  
Published Scientific Articles

**Welcome to the Eros Foundation a biomedical research facility**

**Founded in 1995 by Dr. Millicent Baxter President,**  
Biologist and co-discoverer of pheromones in humans  
and author of:  
[Hormones and your Health: The Smart Woman's Guide to Hormonal and Alternative Therapies for Menopause](#)  
[To Order Click Here](#)

**Eros Science...**  
Dr. Baxter  
**Scholarly Peer-Reviewed Published Eros Science**  
In December 2010, the International Menopause Society journal published Dr. Baxter's invited review of a published paper showing breast cancer seemed to decrease. The 2002 WHI study and subsequent media caused women to stop using hormone replacement therapy.  
Baxter, M, McColl NL, Leiberman, E, Calabrese-Stone E, (2000) Sexual response in women. Obstetrics & Gynecology 95:4(Supplement) April 2000, 19S.  
Baxter M, McColl NL, Leiberman, E (1998) Pheromonal Influences on Sociosexual Behavior. Archives of Sexual Behavior 24:1-13  
Baxter M, Calabrese-Stone E (1998) Women After 40 Years of Age: The Role of Sex Hormones and Pheromones. Disease-A-Month, 44:423-546  
[Breast Cancer in Postmenopausal Women: What is the Real Risk?](#) Dr. Baxter's oral presentation and abstract was accepted for the 65th annual meeting of the American Society of Reproductive Medicine (ASRM) October 2009

**Our Products are shipped in Plain Packages to Protect your Privacy**

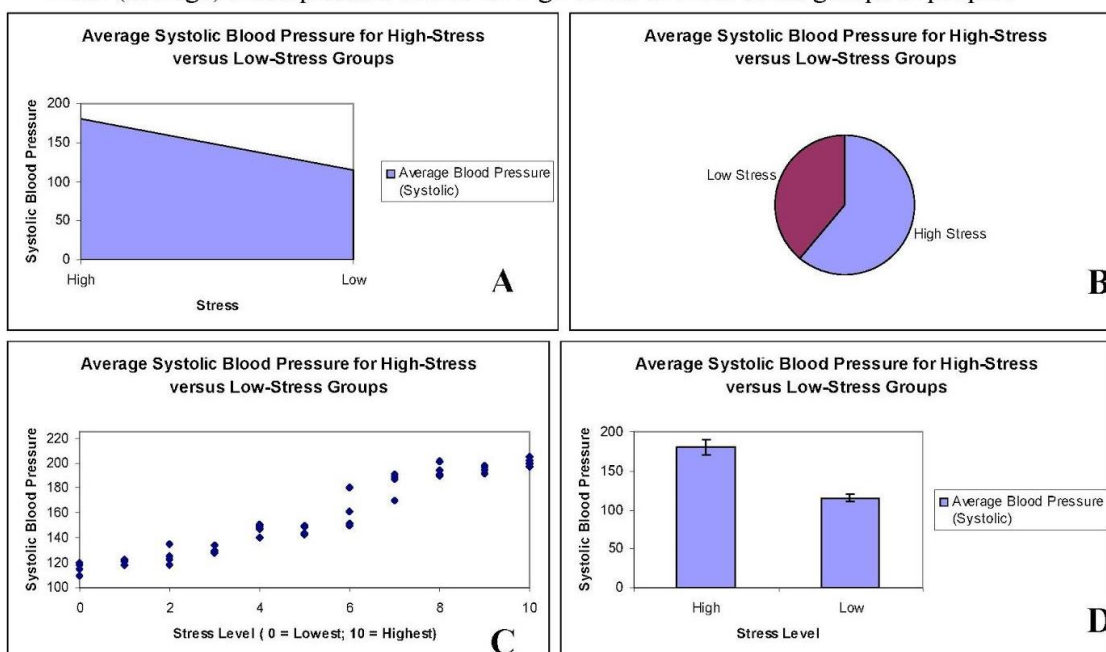
10. For this website (Eros Foundation), which of the following characteristics is **most important** in your confidence that the resource is accurate or not.
- The resource may not be accurate, because appropriate references are not provided.
  - The resource may not be accurate, because the purpose of the site is to advertise a product.
  - The resource is likely accurate, because appropriate references are provided.
  - The resource is likely accurate, because the website's author is reputable.

**Background for questions 11 – 14:** Use the excerpt below (modified from a recent news report on MSNBC.com) for the next few questions.

*“A recent study, following more than 2,500 New Yorkers for 9+ years, found that people who drank diet soda every day had a 61% higher risk of vascular events, including stroke and heart attack, compared to those who avoided diet drinks. For this study, Hannah Gardner’s research team randomly surveyed 2,564 New Yorkers about their eating behaviors, exercise habits, as well as cigarette and alcohol consumption. Participants were also given physical check-ups, including blood pressure measurements and blood tests for cholesterol and other factors that might affect the risk for heart attack and stroke. The increased likelihood of vascular events remained even after Gardener and her colleagues accounted for risk factors, such as smoking, high blood pressure and high cholesterol levels. The researchers found no increased risk among people who drank regular soda.”*

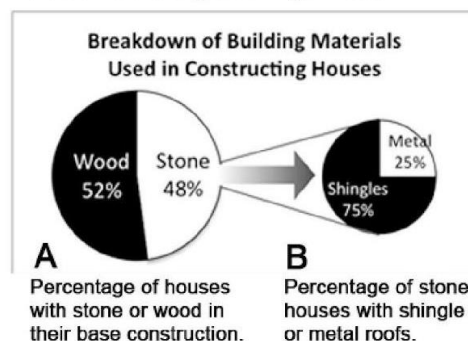
11. The findings of this study suggest that consuming diet soda might lead to increased risk for heart attacks and strokes. From the statements below, identify **additional evidence that supports** this claim:
  - a. Findings from an epidemiological study suggest that NYC residents are 6.8 times more likely to die of vascular-related diseases compared to people living in other U.S. cities.
  - b. Results from an experimental study demonstrated that individuals randomly assigned to consume one diet soda each day were twice as likely to have a stroke compared to those assigned to drink one regular soda each day.
  - c. Animal studies suggest a link between vascular disease and consumption of caramel-containing products (ingredient that gives sodas their dark color).
  - d. Survey results indicate that people who drink one or more diet soda each day smoke more frequently than people who drink no diet soda, leading to increases in vascular events.
12. The excerpt above comes from what type of source of information?
  - a. Primary (Research studies performed, written and then submitted for peer-review to a scientific journal.)
  - b. Secondary (Reviews of several research studies written up as a summary article with references that are submitted to a scientific journal.)
  - c. Tertiary (Media reports, encyclopedia entries or documents published by government agencies.)
  - d. None of the above
13. The lead researcher was quoted as saying, “I think diet soda drinkers need to stay tuned, but I don’t think that anyone should change their behaviors quite yet.” Why didn’t she warn people to stop drinking diet soda right away?
  - a. The results should be replicated with a sample more representative of the U.S. population.
  - b. There may be significant confounds present (alternative explanations for the relationship between diet sodas and vascular disease).
  - c. Subjects were not randomly assigned to treatment and control groups.
  - d. All of the above
14. Which of the following attributes is **not** a strength of the study’s research design?”
  - a. Collecting data from a large sample size.
  - b. Randomly sampling NYC residents.
  - c. Randomly assigning participants to control and experimental groups.
  - d. All of the above.

15. Researchers found that chronically stressed individuals have significantly higher blood pressure compared to individuals with little stress. Which graph would be most appropriate for displaying the mean (average) blood pressure scores for high-stress and low-stress groups of people?



**Background for question 16:** Energy efficiency of houses depends on the construction materials used and how they are suited to different climates. Data was collected about the types of building materials used in house construction (results shown below). Stone houses are more energy efficient, but to determine if that efficiency depends on roof style, data was also collected on the percentage of stone houses that had either shingles or a metal roof.

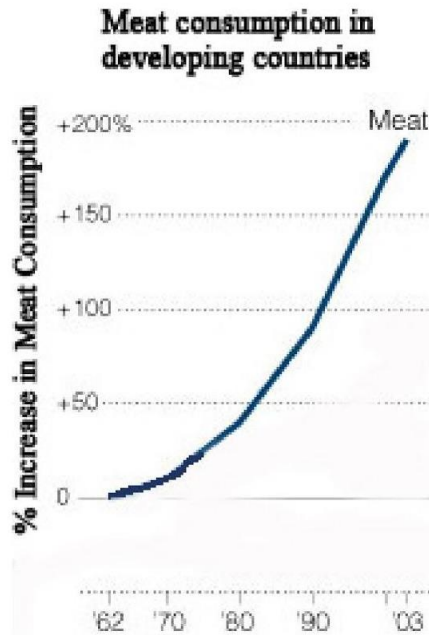
16. What proportion of houses were constructed of a stone base with a shingled roof?
- 25%
  - 36%
  - 48%
  - Cannot be calculated without knowing the original number of survey participants.



17. The **most important** factor influencing you to categorize a research article as trustworthy science is:
- the presence of data or graphs
  - the article was evaluated by unbiased third-party experts
  - the reputation of the researchers
  - the publisher of the article



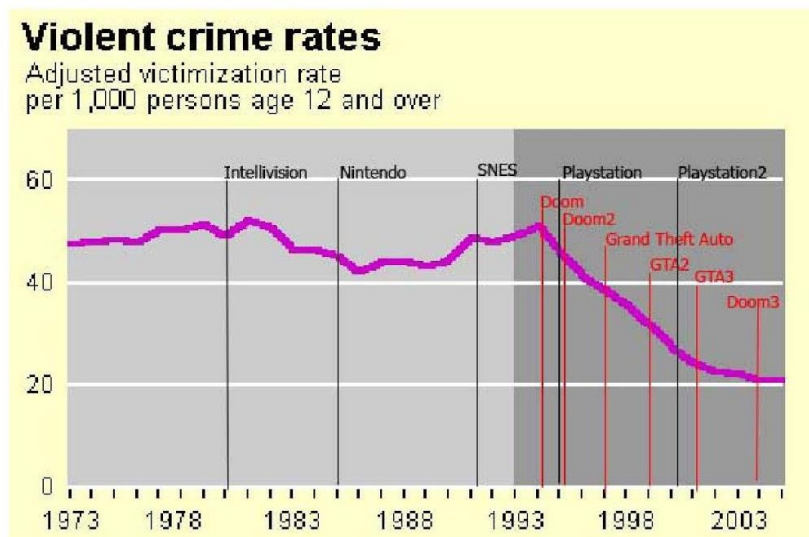
18. Which of the following is the **most accurate** conclusion you can make from the data in this graph<sup>3</sup>?



- a. The largest increase in meat consumption has occurred in the past 20 years.
  - b. Meat consumption has increased at a constant rate over the past 40 years.
  - c. Meat consumption doubles in developing countries every 20 years.
  - d. Meat consumption increases by 50% every 10 years.
19. Two studies estimate the mean caffeine content of an energy drink. Each study uses the same test on a random sample of the energy drink. Study 1 uses 25 bottles, and study 2 uses 100 bottles. Which statement is true?
- a. The estimate of the actual mean caffeine content from each study will be equally uncertain.
  - b. The uncertainty in the estimate of the actual mean caffeine content will be smaller in study 1 than in study 2.
  - c. The uncertainty in the estimate of the actual mean caffeine content will be larger in study 1 than in study 2.
  - d. None of the above
20. A hurricane wiped out 40% of the wild rats in a coastal city. Then, a disease spread through stagnant water killing 20% of the rats that survived the hurricane. What percentage of the original population of rats is left after these 2 events?
- a. 40%
  - b. 48%
  - c. 60%
  - d. Cannot be calculated without knowing the original number of rats.

<sup>3</sup> Modified from Rosenthal, Elizabeth. 2008. As More Eat Meat, a Bid to Cut Emissions. New York Times, December 3, 2008. Accessed June 9, 2011 <http://www.nytimes.com/2008/12/04/science/earth/04meat.html>

**Background for question 21:** A videogame enthusiast argued that playing violent video games (e.g., Doom, Grand Theft Auto) does not cause increases in violent crimes as critics often claim. To support his argument, he presents the graph below. He points out that the rate of violent crimes has decreased dramatically, beginning around the time the first “moderately violent” video game, Doom, was introduced.



21. Considering the information presented in this graph, what is the **most critical flaw** in the blogger's argument?
  - a. Violent crime rates appear to increase slightly after the introduction of the Intellivision and SNES game systems.
  - b. The graph does not show violent crime rates for children under the age of 12, so results are biased.
  - c. The decreasing trend in violent crime rates may be caused by something other than violent video games.
  - d. The graph only shows data up to 2003. More current data are needed.
  
22. Your doctor prescribed you a drug that is brand new. The drug has some significant side effects, so you do some research to determine the effectiveness of the new drug compared to similar drugs on the market. Which of the following sources would provide the **most accurate** information?
  - a. the drug manufacturer's pamphlet/website
  - b. a special feature about the drug on the nightly news
  - c. a research study conducted by outside researchers
  - d. information from a trusted friend who has been taking the drug for six months
  
23. A gene test shows promising results in providing early detection for colon cancer. However, 5% of all test results are falsely positive; that is, results indicate that cancer is present when the patient is, in fact, cancer-free. Given this false positive rate, how many people out of 10,000 would have a false positive result and be alarmed unnecessarily?
  - a. 5
  - b. 35
  - c. 50
  - d. 500

24. Why do researchers use statistics to draw conclusions about their data?
- Researchers usually collect data (information) about everyone/everything in the population.
  - The public is easily persuaded by numbers and statistics.
  - The true answers to researchers' questions can only be revealed through statistical analyses.
  - Researchers are making inferences about a population using estimates from a smaller sample.
25. A researcher hypothesizes that immunizations containing traces of mercury **do not** cause autism in children. Which of the following data provides the **strongest** test of this hypothesis?
- a count of the number of children who were immunized and have autism
  - yearly screening data on autism symptoms for immunized and non-immunized children from birth to age 12
  - mean (average) rate of autism for children born in the United States
  - mean (average) blood mercury concentration in children with autism

**Background for Question 26:** You've been doing research to help your grandmother understand two new drugs for osteoporosis. One publication, *Eurasian Journal of Bone and Joint Medicine*, contains articles with data only showing the effectiveness of one of these new drugs. A pharmaceutical company funded the *Eurasian Journal of Bone and Joint Medicine* production and most advertisements in the journal are for this company's products. In your searches, you find other articles that show the same drug has only limited effectiveness.

26. Pick the **best** answer that would help you decide about the credibility of the *Eurasian Journal of Bone and Joint Medicine*:
- It is not a credible source of scientific research because there were advertisements within the journal.
  - It is a credible source of scientific research because the publication lists reviewers with appropriate credentials who evaluated the quality of the research articles prior to publication.
  - It is not a credible source of scientific research because only studies showing the effectiveness of the company's drugs were included in the journal.
  - It is a credible source of scientific research because the studies published in the journal were later replicated by other researchers.
27. Which of the following actions is a valid scientific course of action?
- A scientific journal rejects a study because the results provide evidence against a widely accepted model.
  - The scientific journal, *Science*, retracts a published article after discovering that the researcher misrepresented the data.
  - A researcher distributes free samples of a new drug that she is developing to patients in need.
  - A senior scientist encourages his graduate student to publish a study containing ground-breaking findings that cannot be verified.

**Background for question 28:** Researchers interested in the relation between River Shrimp (*Macrobrachium*) abundance and pool site elevation, presented the data in the graph below. Interestingly, the researchers also noted that water pools tended to be shallower at higher elevations.

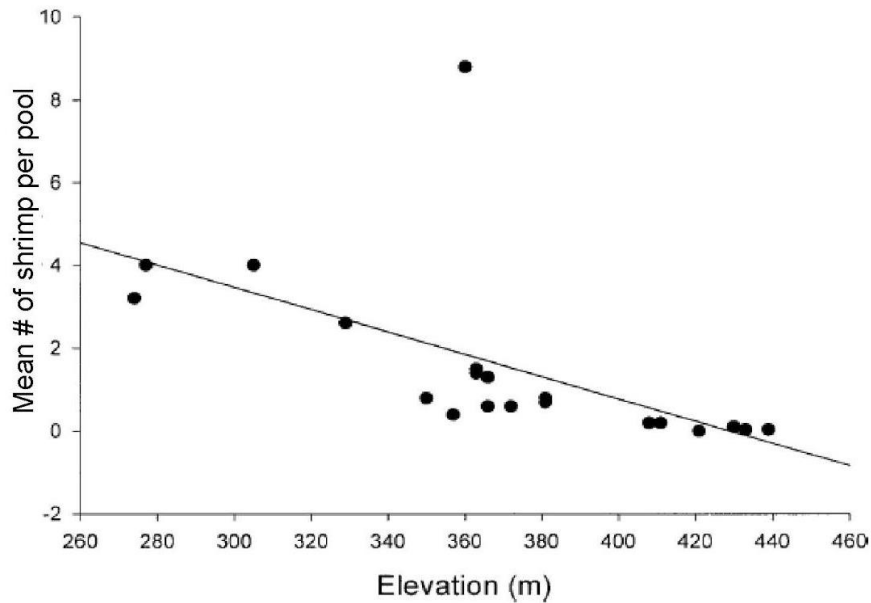


FIG. 3. Relationship between total abundance of *Macrobrachium* (1988–2002) and elevation in Quebrada Prieta.

28. Which of the following is a plausible hypothesis to explain the results presented in the graph?
- There are more water pools at elevations above 340 meters because it rains more frequently in higher elevations.
  - River shrimp are more abundant in lower elevations because pools at these sites tend to be deeper.
  - This graph cannot be interpreted due to an outlying data point.
  - As elevation increases, shrimp abundance increases because they have fewer predators at higher elevations.



<b>SKILL</b>	<b>DESCRIPTION</b>	<b>Question</b>	<b>Correct Answer</b>
<b>1</b> (3 Qs)	Identify a valid scientific argument (e.g., recognizing when scientific evidence supports a hypothesis)	1	B
		8	D
		11	B
<b>2</b> (5 Qs)	Conduct an effective literature search (e.g. Evaluate the validity of sources (e.g., websites, peer reviewed journals) and distinguish between types of sources)	10	B
		12	C
		17	B
		22	C
		26	C
<b>3</b> (3 Qs)	Evaluate the use and misuse of scientific information (e.g. Recognize a valid scientific course of action, distinguish the appropriate use of science to make societal decisions)	5	D
		9	B
		27	B
<b>4</b> (4 Qs)	Understand elements of research design and how they impact scientific findings/conclusions (e.g. identify strengths and weaknesses in research related to bias, sample size, randomization, experimental control)	4	C
		13	D
		14	C
		25	B
<b>5</b> (1Q)	Make a graph	15	D
<b>6</b> (4 Qs)	Read and interpret graphical representations of data	2	C
		6	C
		7	A
		18	A
<b>7</b> (3 Qs)	Solve problems using quantitative skills, including probability and statistics (e.g calculate means, probabilities, percentages, frequencies)	16	B
		20	B
		23	D
<b>8</b> (3 Qs)	Understand and interpret basic statistics (e.g. interpret error bars, understand the need for statistics)	3	B
		19	C
		24	D
<b>9</b> (2 Qs)	Justify inferences, predictions, and conclusions based on quantitative data	21	C
		28	B

## **APPENDIX B: PILOT INTERVIEW SCRIPT**

“Please think aloud as you answer the following question” (See #2 on TOSLS):

- a. “Why do you think that is the correct answer?”
- b. “Why do you think that the other answers are incorrect?”
- c. “Is there anything you recalled from EES 100 that helped you answer this question?”

“Please think aloud as you answer the following question” (See #5 on TOSLS):

- a. “Why do you think that is the correct answer?”
- b. “Why do you think that the other answers are incorrect?”
- c. “Is there anything you recalled from EES 100 that helped you answer this question?”

“Please think aloud as you answer the following question” (See #17 on TOSLS):

- a. “Why do you think that is the correct answer?”
- b. “Why do you think that the other answers are incorrect?”
- c. “Is there anything you recalled from EES 100 that helped you answer this question?”

“Please think aloud as you answer the following question” (See #22 on TOSLS):

- a. “Why do you think that is the correct answer?”
- b. “Why do you think that the other answers are incorrect?”
- c. “Is there anything you recalled from EES 100 that helped you answer this question?”

“Please think aloud as you answer the following question” (See #28 on TOSLS):

- a. “Why do you think that is the correct answer?”
- b. “Why do you think that the other answers are incorrect?”
- c. “Is there anything you recalled from EES 100 that helped you answer this question?”

## APPENDIX C: PRE-INTERVIEW SCRIPT

Review the Consent form, ensure all of the form is understood, and start recording

Ask and record the following demographic data, students do not have to offer this information.

Expected Graduation:\_\_\_\_\_

Major:\_\_\_\_\_

Gender:\_\_\_\_\_

1. Hand student first question:
  - a. "Please read and answer the following question. Feel free to write on the paper."
  - b. "How did you come to this answer?"
2. Hand the student the second question:
  - a. "Please read and answer the following question. Feel free to write on the paper."
  - b. "How did you come to this answer?"
3. Hand student third question:
  - a. "Please read and answer the following question"
  - b. "How did you come to this answer?"
4. Hand student fourth question:
  - a. "Please read and answer the following question"
  - b. "How did you come to this answer?"

#### APPENDIX D: POST-INTERVIEW SCRIPT

Review the Consent form, ensure all of the form is understood, and start recording

Ask and record the following demographic data, students do not have to offer this information.

Expected Graduation: \_\_\_\_\_

Major: \_\_\_\_\_

Gender: \_\_\_\_\_

1. Hand student first question:
  - a. "Please read and answer the following question. Feel free to write on the paper."
  - b. "How did you come to this answer?"
  - c. "Was there anything from EES100 that you did or learned that helped you solve this?"
2. Hand the student the second question:
  - a. "Please read and answer the following question"
  - b. "How did you come to this answer?"
  - c. "Was there anything from EES100 that you did or learned that helped you solve this?"
3. Hand student third question:
  - a. "Please read and answer the following question"
  - b. "How did you come to this answer?"
  - c. "Was there anything from EES100 that you did or learned that helped you solve this?"
4. Hand student fourth question:
  - a. "Please read and answer the following question"
  - b. "How did you come to this answer?"
  - c. "Was there anything from EES100 that you did or learned that helped you solve this?"

Once complete with all questions, ask the following:

1. "Do you remember the Graph Choice Chart from the semester?"

- a. "Did you attend that lecture?"
  - b. "Did you find the activity fun and or helpful?"
- 2. "Do you remember the Summit activity from the semester?"
  - a. "Did you attend that lecture?"
  - b. "Did you find the activity fun and or helpful?"
- 3. "Do you remember the research stakeholder role-play from the semester?"
  - a. "Did you attend that lecture?"
  - b. "Did you find the activity fun and or helpful?"

## **APPENDIX E: DATA SUMMIT- “IS THE ENVIRONMENT OF MAINE PRISTINE?”**

Answer the following questions:

1. Which chart or graph did your group use in class? Include the page number from the pdf that contains all the graphs used in class (available on Blackboard).
2. Based on the figure your group used, what is your answer: Is the environment of Maine pristine?
3. Provide 1 sentence that summarizes whether you think this chart or graph is reliable and relevant to the question. For example, you might say, “This figure showing population growth was derived from census data collected by the U.S. government so I think it is reliable, scientific information”.
4. Based on the graph choice chart (hint: handout in class or slides from the Population Biology lecture), was the appropriate graph or figure used? If yes, list the steps on the graph choice chart that get you to that graph. If no, identify the appropriate graph that should be used and list the steps from the graph choice chart that you used to come to that conclusion.
5. There are often individuals or groups who have vested interest in ongoing research (examples: companies, organizations, manufacturers, etc.). These “stakeholders” can sometimes present data in a way that is advantageous to their interests. What outside factors could have influenced the graphical presentation of the data you chose? Does anyone have a stake in the information that might explain why the graph was presented the way it was?
6. Use the data from your graph to support or refute the following hypothesis (for help, enjoy this video: <https://www.youtube.com/watch?v=HZ9xZHWY0mw>). Hypothesis: If Maine’s natural resources continue to decline in quantity and quality, then Maine is no longer pristine because there has been a negative change in the environment.

## APPENDIX F: DATA SUMMIT- “ARE WE IN THE ANTHROPOCENE?”

Answer the following questions:

1. Choose a graph that, using data, somehow provides evidence to help answer the following question: are we in a new geological time period called the Anthropocene? Bring your graph to class, and attach it to the answers to the following questions.
2. Based on the figure you are using, what is your answer: are we in the Anthropocene? Why or why not? Use the data in your graph to support your claim.
3. Be sure to cite your source, and know where the data come from. Provide 1 sentence that summarizes whether you think this chart or graph is a reliable source and relevant to the question. For example, you might say, “This figure showing population growth was derived from census data collected by the U.S. government so I think it is reliable, scientific information.”
4. Based on the graph choice chart (hint: handout in class or slides from the Population Biology lecture), was the appropriate graph/figure used? If yes, list the steps on the graph choice chart that get you to that graph. If no, identify a more appropriate graph that could be used and list the steps from the graph choice chart that you used to come to that conclusion.
5. There are often individuals or groups who have vested interest in ongoing research (examples: companies, organizations, manufacturers, etc.). These “stakeholders” can sometimes present data in a way that is advantageous to their interests. What outside factors could have influenced the graphical presentation of the data you chose? Does anyone have a stake in the information that might explain why the graph was presented the way it was?
6. Use the data from your graph to support or refute the following hypothesis (for help, enjoy this video: <https://www.youtube.com/watch?v=HZ9xZHWHY0mw>). Hypothesis: Human activities are the predominant driver shaping the earth system, so we are currently in the Anthropocene.

## **APPENDIX G: ROLE PLAY**

### **Scenario:**

The Indian government is considering banning DDT (it is still in use there, and it is still produced in China), which is currently used to control mosquitoes to inhibit the spread of malaria. A meeting of stakeholders will discuss implications and how they should proceed.

### **Roles:**

1. Doctor working for a non-profit organization that provides healthcare to the rural, low-income population locally
2. Representative of a company that farms food in India that is shipped to the U.S.
3. Ecologist who works at a national park in India that is near agricultural land where DDT is used
4. Local spokesperson who represents local Indians that have experienced birth defects due to DDT
5. Resident who has had 3 family members die in areas untreated by DDT

### **Follow-up Think, Pair, Share:**

You are a scientist who has researched the pesticides on food imported to the U.S. from India. You have written your paper, but before publishing an unbiased third-party must evaluate the data, graphs, and paper to ensure the conclusions are sound.

1. Were any of the roles played by the actors an unbiased party? Why or why not?
2. Who would you consider an unbiased 3rd party (or more than one)?

### **Class Instruction:**

1. Have students get into their groups
2. Split the class into 5 sections and assign each section a role
3. Each group in that section will discuss will create arguments for or against banning DDT based on their role- at least 1 person should take notes (7 minutes).



4. Groups from the sections will convene and select a single actor/actress to play the role and give them a script with points of discussion (10 minutes)
5. Have actors/actresses discuss/role play in front of the class, students should take notes on what was discussed (3 minutes)

**Instructional Resource Sheet:**

Role: \_\_\_\_\_

Circle: In favor of banning DDT

Against banning DDT

Points for/against DDT based on role:

Discussions to be had with other roles (Doctor, Ecologist, Company Rep, Spokesperson):

Compromises you would make (if any):

# APPENDIX H: CLASS DATA (MATCHED)

	Overall	Men	Women	EES*	Non-EES**	First-year	Upperclass
	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>
<b>n</b>	44	18	24	12	32	30	14
<b>Mean (% correct)</b>	62	66	58	61	63	61	66
<b>Median (% correct)</b>	62	70	57	57	64	59	71
<b>SD (% correct)</b>	17	18	16	13	19	16	20
<b>Variance (% correct)</b>	3	3	2	2	3	2	4
<b>Min (% correct)</b>	25	25	32	46	25	32	25
<b>Max (% correct)</b>	96	93	96	82	96	93	96
	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>
<b>n</b>	44	18	24	12	32	30	14
<b>Mean (% correct)</b>	63	62	64	68	61	63	63
<b>Median (% correct)</b>	64	68	64	64	66	64	61
<b>SD (% correct)</b>	19	22	14	14	21	18	22
<b>Variance (% correct)</b>	3	5	2	2	5	3	5
<b>Min (% correct)</b>	21	21	25	50	21	21	25
<b>Max (% correct)</b>	96	89	93	96	93	96	93

\*Ecology and Environmental Science majors

\*\*Majors other than EES (Non-EES)

# APPENDIX I: CLASS DATA (UNMATCHED)

	Overall	Men	Women	EES*	Non-EES**	First-year	Upperclass
	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>	<i>Pre</i>
<b>n</b>	88	42	42	26	62	53	35
<b>Mean (% correct)</b>	58	59	56	54	60	56	61
<b>Median (% correct)</b>	57	59	54	52	61	54	61
<b>SD (% correct)</b>	20	22	18	18	20	20	20
<b>Variance (% correct)</b>	4	5	3	3	4	4	4
<b>Min (% correct)</b>	18	18	18	18	18	18	25
<b>Max (% correct)</b>	96	93	96	82	96	93	96
	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>	<i>Post</i>
<b>n</b>	56	25	27	14	42	37	18
<b>Mean (% correct)</b>	60	57	64	69	57	60	62
<b>Median (% correct)</b>	61	61	64	64	59	61	64
<b>SD (% correct)</b>	20	22	15	14	20	18	22
<b>Variance (% correct)</b>	4	5	2	2	4	3	5
<b>Min (% correct)</b>	21	21	29	50	21	21	25
<b>Max (% correct)</b>	96	89	93	96	93	96	93

\*Ecology and Environmental Science majors

\*\*Majors other than EES (Non-EES)

## **BIOGRAPHY OF THE AUTHOR**

Molly Picillo was born in Newburyport Massachusetts on August 31<sup>st</sup>, 1994. She was raised there and graduated from Newburyport High School in 2012. She attended the University of Maine and graduated in 2016 with a Bachelor's degree in Zoology and Wildlife Ecology. After field ecology experiences in Wisconsin and Costa Rica, she returned to Maine and entered the RiSE teaching graduate program at The University of Maine in the Spring of 2017. Upon receiving her degree, Molly will be pursuing field work in ecology where she can utilize her educational background in public outreach on topics of human dimensions, conservation, and preservation. Molly is a candidate for the Master of Science in Teaching degree from the University of Maine in August 2018.